



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학박사학위논문

# 좌측절단 자료를 이용한 수리가능 시스템의 신뢰도 추세분석

**A reliability trend analysis for heterogeneous  
repairable systems with left truncated data**

2018년 2월

서울대학교 대학원  
산업공학과  
나 일 용



# 좌측절단 자료를 이용한 수리가능 시스템의 신뢰도 추세분석

지도교수 장 우 진

이 논문을 공학박사 학위논문으로 제출함  
2017년 12월

서울대학교 대학원  
산업공학과  
나 일 용

나일용의 박사 학위논문을 인준함  
2018년 1월

위 원 장 박 용 태 (인)

부위원장 장 우 진 (인)

위 원 조 성 준 (인)

위 원 이 재 욱 (인)

위 원 배 석 주 (인)





# 초 록

## 좌측절단 자료를 이용한 수리가능 시스템의 신뢰도 추세분석

서울대학교 대학원  
산업공학과  
나 일 용

첨단 시스템의 사용 품질을 꾸준히 유지하는 것은 어려운 일이다. 제품의 복잡함으로 인해, 고장 발생 가능성은 높은 반면 탐지와 복구는 난해하기 때문이다. 따라서 제품의 복잡도 증가에 따라 고장의 발생을 최소화하고 발생한 고장을 최단시간에 최소한의 노력으로 복구할 수 있는 제품 설계와 운영유지 정책이 그 무엇보다 중요한 요소로 여겨지고 있다. 위 활동에 가장 큰 영향을 주는 요소가 바로 “신뢰성”이다. 사전적 의미의 신뢰성은 “주어진 조건에서 정해진 시간동안 고장 없이 사용할 수 있는 확률”로 정의되며, 핵심 키워드가 바로 “고장”이기 때문이다.

고장을 최소화 하기 위해, 개발자는 고 신뢰성 부품을 사용하거나 부품의 이중화, 단순화 된 구성품의 사용 등 다양한 방법을 사용하곤 한다. 만약 고장을 최소화 할 수 없을 경우, 정비를 좀 더 쉽게하여 복구시간을 단축시키기도 한다. 다만, 개발이 완료된 후에는 추가적인 개선이 어렵기 때문에, 개선보다는 현상 유지를 위한 활동에 중점을 두게 된다. 즉, 장비 상태를 모니터링하여, “고장의 추이”가 증가되는 지, 감소되는 지, 일정 수준으로 유지되고 있는 지를 관찰하게 되는 것이다.

수리가능 시스템에서는 고장의 추이를 관측하는 모형으로 Power Law

모형을 보편적으로 사용해 왔다. Power Law 모형은 와이블 함수 형태의 강도함수를 사용하는 NHPP모형으로, 모형의 단순함과 사용의 편리함으로, 70년대 Crow가 제안한 이래 현재까지 무기체계 등 수리가능 시스템과 관련된 분야에서 꾸준히 사용 중이다. 그러나 이 같은 단순함과 편리함에도 불구하고, 모형의 사용을 위해서는 장비운영 초기부터 꾸준히 수집된 자료를 사용해야 한다는 제약이 있다. 특히 수명주기(Life Cycle)가 짧아 고장자료를 손쉽게 추적하고 유지할 수 있는 상용품에 비해, 수명주기가 보통 10년 이상인 군용장비는 자료의 품질 문제로 모형의 사용에 제한이 발생하기도 한다. 국내 역시 장비의 수명주기 동안 자료관리 인원이나 기록시스템의 잦은 변경으로 인해 자료 손실이 빈번하게 발생되어 왔다. 이에 따라, 자료의 양은 많으나 대부분 좌측절단된 자료이기 때문에, 실제 활용할 수 있는 분석 모형이 제한되고, 특히 신뢰도 추세분석과 같은 분석에는 사용하기 어려운 측면이 있다.

본 연구는 이러한 좌측절단 상황에서도 신뢰도 추세분석을 할 수 있도록 Power Law 모형을 변형한 실용적인 모형 제시를 목적으로 하고 있다. 좌측절단 상황을 Power Law 모형에 반영하기 위해, 손실을 보상하는 방법으로 E-M 알고리즘의 사용을 제안하고 있다. E-M 알고리즘은 자료 손실을 보상하기 위한 일반적 방법으로, 본 논문에서와 같이 확률모형의 모수추정에 ML(Maximum Likelihood)을 사용하는 경우 최적화된 방법이라고 할 수 있다. 다만, 전역 최적화를 보장할 수 없는 단점이 있어 이를 보완하기 위해 경험적 방법에 의한 모수 추정치를 초기값으로 하는 방안을 제안하였다. 또한, 비록 동일한 장비라 하더라도 사용 방법과 사용자, 사용 환경 등에 따라 각기 상이한 고장추세를 나타낼 수 있기 때문에 이를 고려할 수 있는 확장된 모형을 별도로 제시하였다. 전자를 “기본모형”, 후자를 “확장모형”으로 칭하며, 확장모형은 관측 가능한 이질요소인 공

변량과 Frailty를 동시에 고려할 수 있도록 제안하였다.

제안된 모형의 효용성과 활용 범위를 확인하기 위해 다양한 자료손실 조건과 Power Law 모수 및 Frailty 수준 조합에 대해 모의실험을 실시하였고, 제안한 모형의 활용범위와 오차수준 등을 확인하였다. 기본모형의 경우 자료손실 수준 10%이하에서 유효한 것으로 확인되었다. 또한 확장 모형이 기본모형에 비해 자료손실 수준에 관계 없이 좀 더 균일한 추정 오차를 나타내고 있는 것으로 확인되었다. 마지막으로, 제안된 두 모형을 국내에서 사용 중인 전차의 야전 고장자료에 적용하여 해당장비의 신뢰도 수준을 산출하였고, 이를 활용한 간단한 정비정책 결정 사례를 제시함으로써, 실제 정비정책 결정에 충분히 활용할 수 있음을 나타내고자 하였다.

본 연구는 좌측절단된 자료를 효과적으로 처리하고자 하는 문제 의식에서 시작되었고, 일정 범위 내에서 유효한 결과를 산출하는 모형을 제시할 수 있었다. 다만, 대표본 가정의 빈번한 사용, 수리적 모형의 제한적 제시, 한정적 범위 내에서의 민감도 분석 수행 등은 추후 연구를 통해 점진적으로 보완되어야 할 것으로 판단된다.

**주요어 :** 좌측절단, 신뢰도 추세, 비동질성, Frailty, 공변량, 기본모형, 확장모형

**학번 :** 2011-30312



# 목 차

<b>I. 서론</b>	<b>1</b>
1.1 연구 배경 및 필요성	1
1.2 연구 범위	3
1.3 논문의 구성	4
<b>II. 배경 이론</b>	<b>7</b>
2.1 분석 대상자료의 특성	7
2.2 신뢰도 추세분석의 목적과 영향	11
2.3 관련 연구	17
2.3.1 연구현황	17
2.3.2 기존연구의 한계점	22
2.3.3 시사점과 연구방안 수립	23
<b>III. 신뢰도 추세분석 기본모형</b>	<b>25</b>
3.1 좌측절단을 고려한 모수추정 방안	27
3.2 모수추정	33
3.2.1 추정의 실효성을 높일 수 있는 초기값 추정	34
3.2.2 E-M 알고리즘의 활용	35
3.2.3 신뢰구간 추정	36
3.3 제안된 모형의 성능 확인	39
3.3.1 모형의 효용성과 민감도	40
3.3.2 신뢰구간의 정확성	46

<b>IV. 이질요소를 고려한 모형의 확장</b>	<b>49</b>
4.1 관측 불가능한 이질요소(Frailty) 반영	50
4.1.1 강도함수의 모수추정	50
4.1.2 모수의 분산과 Frailty의 크기 추정	55
4.2 관측 가능한 이질요소(공변량) 추가 반영	59
4.2.1 공변량 추가모형의 모수추정	59
4.2.2 Frailty의 크기 추정	64
4.3 확장 모형의 성능 확인	65
4.3.1 모형의 민감도와 효용성	66
4.3.2 기본모형과의 성능 비교	78
<b>V. 실 자료를 활용한 실증분석</b>	<b>80</b>
5.1 분석에 사용된 자료의 특성	80
5.2 기본모형 적용결과	83
5.3 확장모형 적용결과	87
5.4 분석 결과의 활용	90
<b>VI. 결론 및 추후연구</b>	<b>94</b>
6.1 연구 결과	94
6.2 한계점 및 추후 연구	97
<b>Abstract</b>	<b>122</b>

# 표 목 차

표 1.	좌측절단된 다수장비의 고장자료 형태 . . . . .	12
표 2.	기본모형의 성능확인을 위한 실험조건 . . . . .	40
표 3.	기본모형의 손실비율에 따른 추정효율( $\overline{NRMSE}$ ) . . . . .	44
표 4.	기본모형의 모수추정 사례 (5% 손실수준) . . . . .	46
표 5.	기본모형의 신뢰구간 산식의 타당성 확인결과 . . . . .	47
표 6.	Frailty 추가모형의 모의실험 결과( $\overline{NRMSE}_\alpha$ ) . . . . .	75
표 7.	Frailty 추가모형의 모의실험 결과( $\overline{NRMSE}_\beta$ ) . . . . .	75
표 8.	기본모형과 확장모형의 모수추정 결과 비교 . . . . .	76
표 9.	공변량 확장모형의 손실비율 별 공변량, $\eta_0$ 추정결과 . . . . .	77
표 10.	K계열 전차의 아전 수집자료 . . . . .	81
표 11.	실 자료의 NHPP여부 확인 결과 . . . . .	83
표 12.	실 자료의 모수추정 결과 . . . . .	83
표 13.	확장모형의 모수추정 결과(Case II.) . . . . .	88



# 그림 목 차

그림 1.	연구수행 개념 . . . . .	4
그림 2.	논문의 구성과 연구 흐름 . . . . .	6
그림 3.	수명주기 비교 . . . . .	8
그림 4.	단일장비의 정상적인 고장자료 기록 . . . . .	9
그림 5.	좌측절단이 발생된 단일장비의 불완전한 고장자료 . . . . .	10
그림 6.	다수 장비의 좌측절단 고장자료 형태 . . . . .	11
그림 7.	다수 장비의 좌측절단 고장자료 생성개념 . . . . .	12
그림 8.	미 육군의 신뢰도 성장관리 계획 사례(Handbook (1981)) . . . . .	14
그림 9.	개선효과의 측정 개념 . . . . .	15
그림 10.	PEXP Model 개념, ( $\delta > 1$ 일 때, Arab et al. (2012)) . . . . .	19
그림 11.	기본모형의 모수추정 절차 . . . . .	26
그림 12.	경험적 방법을 활용한 초기값 추정 . . . . .	35
그림 13.	모의실험 결과 NRMSE(MR : 0.01, 0.02, 0.05) . . . . .	42
그림 14.	모의실험 결과 NRMSE(MR : 0.1, 0.2, 0.3) . . . . .	43
그림 15.	기본모형의 손실비율에 따른 추정효율( $\overline{NRMSE}$ ) . . . . .	45
그림 16.	$\widehat{MTBF}(t)$ 와 신뢰구간 . . . . .	47
그림 17.	$\widehat{m}(t)$ 와 신뢰구간 . . . . .	48
그림 18.	Frailty를 고려한 확장모형의 모수추정 절차 . . . . .	55
그림 19.	공변량을 추가 고려한 확장모형의 모수추정 절차 . . . . .	63
그림 20.	모의실험 결과 NRMSE( $\delta = 0.1, MR = 0.01, 0.02, 0.05$ ) . . . . .	67
그림 21.	모의실험 결과 NRMSE( $\delta = 0.1, MR = 0.1, 0.2, 0.3$ ) . . . . .	68
그림 22.	모의실험 결과 NRMSE( $\delta = 0.3, MR = 0.01, 0.02, 0.05$ ) . . . . .	69

그림 23.	모의실험 결과 $\text{NRMSE}(\delta = 0.3, MR = 0.1, 0.2, 0.3)$ .	70
그림 24.	모의실험 결과 $\text{NRMSE}(\delta = 0.5, MR = 0.01, 0.02, 0.05)$	71
그림 25.	모의실험결과 $\text{NRMSE}(\delta = 0.5, MR = 0.1, 0.2, 0.3)$ . .	72
그림 26.	모의실험 결과 $\text{NRMSE}(\delta = 0.7, MR = 0.01, 0.02, 0.05)$	73
그림 27.	모의실험 결과 $\text{NRMSE}(\delta = 0.7, MR = 0.1, 0.2, 0.3)$ .	74
그림 28.	기본모형과 확장모형의 $\text{NRMSE}$ 비교( $\delta = 0.1$ ) . . . .	79
그림 29.	Case I, II.의 $\widehat{m(t)}$ 와 신뢰구간 비교 . . . . .	84
그림 30.	Case I. 자료와 Case II.의 $\widehat{m(t)}$ 비교 . . . . .	85
그림 31.	Case II. 분석 결과와 Case II. 자료 비교 . . . . .	87
그림 32.	기본모형과 확장모형의 모수추정 결과 비교 . . . . .	89

## Notation

$x_{ji}$	: $j^{th}$ 장비의 $i^{th}$ 고장 발생 시점
$s_j$	: $j^{th}$ 장비의 관측이 시작된 시점
$t_j$	: $j^{th}$ 장비의 관측이 종료된 시점
$r_j$	: $j^{th}$ 장비의 최초 관측된 고장의 order
$\lambda(t)$	: NHPP 모형의 강도함수(Intensity Function)
$m(t)$	: $t$ 시점의 누적 고장 수
$MTBF(t)$	: $t$ 시점의 평균 고장간 시간
$\Lambda[t_0, t_1]$	: $[t_0, t_1]$ 구간에서 발생된 누적 고장 수
$Z$	: Frailty의 크기
$z_j$	: $j^{th}$ 장비의 Frailty
$\delta$	: Frailty의 규모( $\text{Var}(Z)$ )
$\eta_0$	: 개별장비의 공변량 요소
$Y$	: 공변량 유무를 결정하는 지시변수(Indicator variable)
$C_1$	: 고장 정비비용
$C_2$	: 창정비 비용
$T_{ov}$	: 최적 창정비 시점(or 주기)
$TC(T_{ov})$	: 창정비 주기동안 발생된 총 비용
$C(T_{ov})$	: 창정비 주기동안 발생된 평균 정비비용
$IMC(t)$	: $t$ 시점에서의 순간 정비비용

# 제 1 장

## 서론

### 1.1 연구 배경 및 필요성

효과적인 시스템이란 주어진 목적에 충분한 성능을 발휘될 수 있는 장비이다. 충분한 성능이라 함은 단순히 기계·전자적 특성 외에도, 사용 기간 중 고장 없이 성능이 지속적으로 유지되는 것을 의미한다(Bell and Bearden (2014)). 이를 나타내는 장비 고유의 특성을 “신뢰성”이라 한다. 신뢰성은 사용 중 발생하는 정비 빈도나 유지비용과도 밀접한 관련이 있어 제품 개발 시 초기단계부터 폐기시점까지 지속적으로 관리 되어야 할 요소로 인식되어 왔다. 이와 같은 신뢰성을 지표화 한 것을 “신뢰도”라 하며, 대표적인 신뢰도 지표로 시간 당 발생하는 고장수를 의미하는 “고장률(Failure rate)” 또는 고장과 고장 사이의 평균 시간인 “평균 고장 간 시간(MTBF: Mean Time Between Failure)”등이 사용되곤 한다(O’connor et al. (2012)). 신뢰성은 장비 고유 특성이므로, 동일한 장비를 같은 방식으로 같은 환경에서 개조나 개선 없이 사용할 경우 시간에 따라 항상 같은 신뢰도를 나타내야 한다. 그러나 사용에 따른 열화, 사용 기간 중에 발생하는 다양한 개선활동 등의 영향으로 시간에 따른 신뢰도는 일정하지 않은 것이 일반적이다(Wang and Coit (2005)). 이 때문에, 수행 중인 다양한 신뢰성 개선활동의 효과를 확인하거나 현재 장비의 신뢰성 수준을 파악하여 미래의 정비 관련 의사결정을 지원하기 위해 장비의 신뢰도 변화 추세를 분석하곤 한다. 이 같은 분석을 위해 확률모형, 시계열, 회귀분석, Cox모형

등 다양한 방법이 사용되어 왔으며, 이 중 NHPP 기반의 확률모형인 Power Law 과정을 분석의 단순함과 해석의 편이성 때문에 가장 많이 사용 중이다. 미국 등 선진국에서는 위 방법을 무기체계 등의 제품개발 절차에 반영하여 "신뢰성 성장 관리(Reliability Growth Management)"기법으로 명명(Handbook (1981))하고, \$185M 이상 무기체계 개발사업의 관리에 필수적으로 사용하고 있다(Ellner and Herbert (2013)). 최근 국내에서도 군용장비의 복잡화, 다양화에 따라 과거보다 운영 유지비의 절감 요구가 증가되고 있으며, 해법으로 정량적 분석 결과를 바탕으로 한 신뢰성 관리를 시도 중이다.

개발 중인 장비의 경우 시험을 수행하는 장비도 적고 이에 따라 고장이 많이 발생되지 않는다. 또한, 작은 고장이라도 시험 중 고장이 발생하는 현상 자체를 심각하게 인식 때문에, 시험 초기단계부터 엄격하게 자료 관리를 하는 것이 일반적이다. 이에 따라 손실이 거의 없는 상태의 자료를 분석에 사용할 수 있다. 이와 달리 개발 종료 후 실제 사용 중에는 개체 수 증가에 따라 장비 군(Group) 전체적으로 고장이 급증하며 다양한 요인으로 자료도 간혹 손실되곤 한다. 특히 군용장비의 경우 상용품과 달리 10년 이상 매우 장기간 사용되는 것이 보통이기 때문에 기록시스템의 변화로 대량의 자료손실이 발생하는 경우도 있다. 고장을 기록하는 기록시스템의 변화가 대표적인 사례이다. 기록 시스템을 변경하는 과정에서 과거자료 역시 이관은 수행되지만 수기 기록의 전산화, 신-구 시스템 간 자료형식 변환 등과 관련된 예산과 시간의 절감을 위해 일부 자료를 이관에서 배제하거나 원천자료가 아닌 단순화 된 자료로 이관하는 경우가 있다. 실제로 현재 육군의 정비활동을 기록하는 시스템은 후자의 방법으로 기존자료를 이관하였다. 이에 따라, 과거부터 운용중인 장비의 고장실적이 현재 기록되는 자료와 다른 단순화 된 형태로 관리되게 되었다. 그

결과, 새로운 기록시스템 이전부터 운용 중인 장비의 고장 실적은 좌측절단(Left truncated)된 형태의 자료가 된다.

국방 분야에서는 신뢰도 추세분석에 Power Law 과정을 가장 많이 사용해 왔고(Crow (1974); Wayne and Modarres (2015)), 추후 타 장비 분석 결과와의 비교분석 등 활용도를 고려 시 Power Law 과정 사용이 가장 유리하다. 그러나 Power Law 과정에 대해 일반적인 관측중단 상황인 Type I, II 절단을 고려하는 모형, 일부 자료누락을 고려하는 모형 등(Sun et al. (2005); Crow and Basu (1988); Little and Rubin (2014))은 쉽게 찾을 수 있었으나, 위와 같은 좌측절단 상황을 처리할 수 있는 연구는 현재까지 거의 존재하지 않음을 알게 되었다. 물론 Yu et al. (2008)과 유사한 연구가 일부 존재하였으나, 분석 대상 자료와 정확히 일치하지 않고, 유실된 자료의 개수를 알아야 하는 등 실효성이 낮아 활용이 곤란하였다. 이에 Power Law 과정의 기본 특성과 형태를 유지하며 좌측절단 상황을 효과적으로 처리할 수 있는 수정된 모형의 연구 필요성을 인식하게 되었다.

## 1.2 연구 범위

본 연구의 최종 목적은 그림 1.과 같이 다수의 장비에서 수집된 좌측절단 된 고장자료 만으로, 정상자료를 사용하였을 때와 유사한 결과를 산출 할 수 있는 분석모형을 제안하는 것이다. 기반 모형은 활용성을 고려하여 Power Law 형태의 강도(Intensity)함수를 갖는 NHPP 모형을 사용하였다. 이를 통해 기존에 신뢰도 성장관리 활동을 통해 분석한 다양한 결과와 비교분석 할 수 있는 호환성 있는 추정 결과를 산출할 수 있게 된다. 이 과정에서 좌측절단에 의해 손실된 자료를 보상하는 방법으로 ML 추정에서 일반적으로 사용되는 기댓값 최대화(Expectation Maximization,

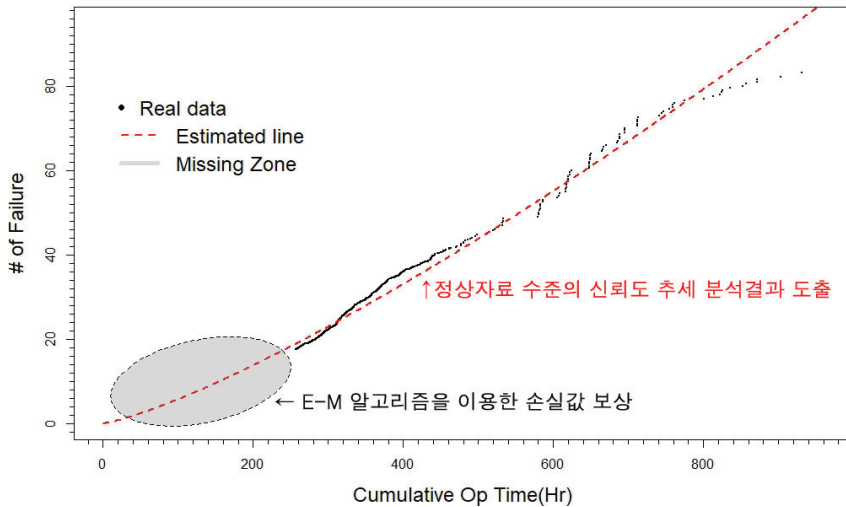


그림 1: 연구수행 개념

이하 E-M) 알고리즘의 사용을 제안하려 한다. 이 과정에서, 전역 최적화를 보장하지 못하는 단점을 보완하고자 누적 고장수의 추세를 도식적으로 분석하여 E-M 알고리즘의 초기값으로 사용하는 방안을 추가로 제안하게 된다. 아울러, 분석 대상이 다수 장비임을 고려하여 장비 간 발생할 수 있는 유·무형의 차이를 고려하기 위한 보완책도 제시하려 한다. 이는 모형의 현실성을 확보하기 위한 것이며, 실제로 동일한 장비임에도 운영 환경이나 사용빈도 등이 장비 고장특성에 영향을 줄 수 있음이 Asfaw and Lindqvist (2015); Slimacek and Lindqvist (2016, 2017)을 통해 알려져 있기 때문이기도 하다.

### 1.3 논문의 구성

본 논문은 총 6개의 장으로 구성된다.

2 장에서는 다수 장비에서 수집된 좌측절단 고장 자료의 형태와 특

이점을 설명한다. 또한 신뢰도 추세분석의 목적, 필요성과 기존에 수행된 관련 연구 결과들을 방법론 별로 소개한다. 마지막 부분에는 기존 연구 결과 중 제한적나 활용 가능한 연구결과를 식별하고, 연구 목표의 달성을 위해 수정반영해야 할 사항 등 기존 연구결과의 시사점들을 정리한다.

3 장에서는 Power Law 기반의 NHPP모형의 모수를 좌측절단 자료를 이용하여 추정하는 실질적인 방법을 제시한다. 추정의 실효성을 높일 수 있는 초기값 추정 방법과 함께, E-M 알고리즘을 활용하기 위한 모수 추정식과 산출 절차를 단계별로 설명하며, 모의실험 결과를 통해 제안한 방법의 타당성, 오차범위 및 유효범위 등을 설명한다.

4 장에서는 연구결과의 실효성을 높이기 위해 다수 장비에서 수집된 자료에서 나타날 수 있는 장비별 비동질적 특성(Frailty와 공변량(Covariate))을 모형에 추가하는 방안을 제시한다. 이 역시 기본모형과 유사하게 모의실험으로 제안된 방법의 적절성을 확인한다.

5 장에서는 3장과 4장에서 제안한 방법을 실제 국내 전차의 운용중에 수집된 고장자료에 적용하여 대상 장비의 신뢰도 변동 특성을 분석하고, 정비정책 결정에 활용하는 사례를 제시한다. 이를 통해 앞서 제안된 모형들이 실 자료 분석에도 실효성 있게 사용할 수 있음을 보여주게 된다.

마지막으로 6 장은 연구의 주요 성과들을 정리하고 향후 연구 방향에 대해 다룬다. 이상의 논문의 구성과 연구 흐름은 그림 2.와 같이 요약된다.





그림 2: 논문의 구성과 연구 흐름

## 제 2 장

## 배경 이론

### 2.1 분석 대상자료의 특성

본 논문의 분석 대상은 실제 운용 중 발생한 고장자료이다. 위와 같은 자료는 실험실 환경에서 수집된 자료와 다음과 같은 차별점을 갖는다. 우선 다수의 장비에서 발생한 자료이다. 성능시험이나 수명시험은 통상 제한된 시간 내에 정해진 예산 범위 내에서 수행되기 때문에 몇 개의 구성품이나 장비를 활용하는 경우가 대부분이다. 이와 달리, 실제 운용 중에는 관측 종료시점까지 사용중인 모든 장비가 관찰 대상이 된다. 즉, 관찰 가능한 장비 수의 큰 증가가 이루어진다. 이 때문에 실험실 환경에서 확인하기 어려운 저 빈도 고장까지도 통계적 분석이 가능한 수준까지 수집 가능해진다. 이러한 이유로, 개발이 종료된 이후에도 지속적인 야전자료 수집·분석은 꾸준히 강조되고 있다. 분석 가능한 자료의 증가라는 장점이 있는 반면, 엄격히 통제되지 않은 상황에서 발생하는 자료라는 단점도 존재한다. 실험실 환경에서는 실험과 관계없는 외부 요인을 엄격하게 통제하는 변인통제가 용이하나 야전자료의 경우 이 과정이 사실상 불가능하다. 대상물의 사용자가 불특정 다수이며, 기본적인 장비 사용법은 동일하나 주변 환경이나 사용빈도 등이 약간씩 상이한 경우가 많다. 또한, 고장과 관계없는 단순 소모품의 보급과 같은 노이즈가 자료에 포함되는 경우가 많다. 이러한 이유로 야전자료는 분석 이전에 유사한 특성을 갖는 자료들만을 선별하고 일정한 기준에 따라 자료를 정제하는 과정도 요구된다. 이렇듯

장점과 단점을 모두 보유 하였지만, 야전 자료 활용의 가장 큰 문제점은 자료의 유실문제이다. 그림 3.과 같이 흔히 사용되는 휴대폰, 가전제품 등 공산품의 경우 대부분 수명주기가 3년 내외이고, 자동차의 경우에도 길어야 10년 정도 사용되는 반면, 전차나 장갑차 같은 군용장비의 경우 짧게는 10년에서 길게는 30년까지 사용되는 경우가 대부분이다. 물론 사용 중에 지속적인 개조나 개선을 수행하는 사례가 많으나, 장비 자체의 수명주기는 공산품에 비해 매우 긴 특성을 갖는다. 국내에서 80년대 초에 개발된 K계열 전차나 자주포가 아직도 주력장비로 사용하는 것이나 미국이 70년대 도입된 F15 전투기, M1 전차를 현재까지도 사용하고 있는 것이 대표적 사례라 할 수 있다.

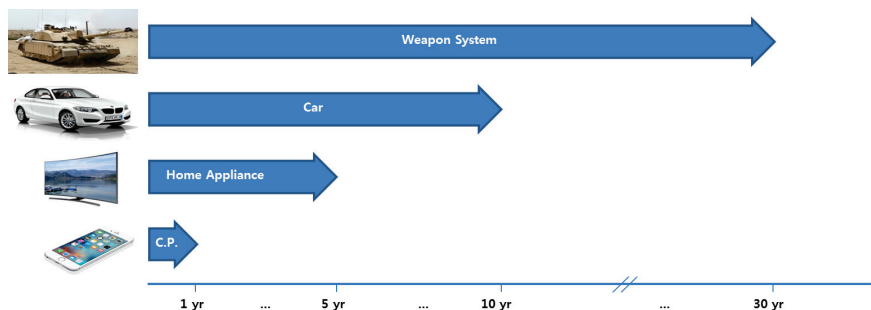


그림 3: 수명주기 비교

이처럼 긴 수명주기 동안 그림 4.와 같이 분석을 위한 정확히 자료가 수집되고 잘 관리되는 것은 매우 어려운 일이다. 적시에 기록되어야 하는 것은 물론 긴 시간동안 손실 없이 관리되어야 한다. 기록의 적시성은 기술 진보에 따라 기록시스템의 고도화를 통해 달성할 수 있다. 그러나 자료의 유실은 과거자료의 중요성에 대한 인식 부족과 자료이관에 따른 예산부족 문제로 빈번하게 발생되어 왔다. 즉, 장비 운영기간 중 발와생되는 기록시스템의 변화가 과거자료의 관리에 큰 저해 요인이 되고 있다. 국내에서는

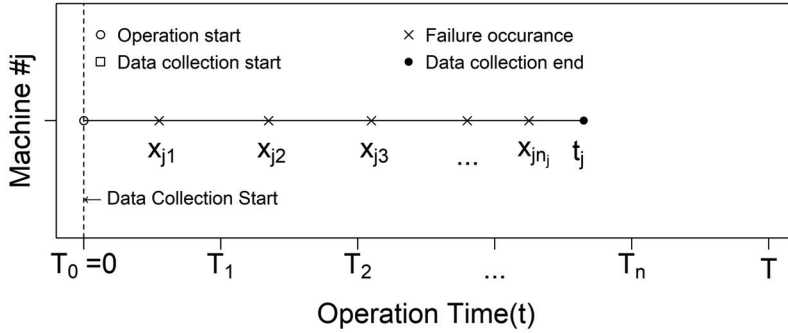


그림 4: 단일장비의 정상적인 고장자료 기록

2000년대 후반부터 야전 고장자료의 수집과 관리를 매우 중요한 것으로 인식하여 장비별 고장발생 시점, 정비시간, 정비 중 사용된 수리부속품 종류와 수량, 수리부속품을 확보하기까지 소요된 시간 등 고장과 관련된 다양한 정보들을 관리할 수 있는 “장비정비정보체계”라는 기록 시스템을 운영 중이다. 이를 통해 보다 신뢰할 수 있는 분석 자료를 확보할 수 있게 되었다. 그러나 과거에는 고장 시점보다는 고장의 회복을 위한 정비 행위만을 중시하였고 야전의 열악한 실태를 고려하여 전산화된 기록보다는 수기 기록물로 자료를 관리하는 경우가 많았다. 이 때문에, 기록시스템이 변경될 때마다, 방대한 자료의 이관에 어려움이 있어 왔다. 장비정비정보 체계의 경우, 대부분의 과거자료가 이관은 되었으나, 현재와 같은 형식이 아닌 요약된 형태로 이관되었다. 특히 고장횟수 정보는 남아있지 않으며, 장비정비정보체계 사용 이전까지 장비 운용량 정보만 남아있는 실정이다.

그림 5.는 현재 신뢰도 추이 분석의 대상이 되는 원천자료의 형태를 장비 1대(Machine #j)를 기준으로 도식화 한 것이다.  $x_{ji}$ 는 위 장비에서 발생된  $i^{th}$  고장을 의미하며, 총  $t_j$ 까지 관측된 상태를 나타낸다. 또한,  $n_j$

는 해당 기간 중 발생된 장비  $j$ 의 총 고장수를,  $s_j$ 는 실제 관측이 시작된 시점을 의미한다. 통상 신뢰도는 장비가 실제 가동된 상태만을 기준으로

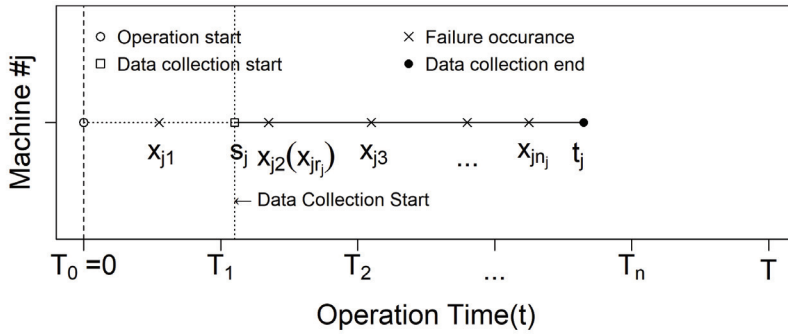


그림 5: 좌측절단이 발생된 단일장비의 불완전한 고장자료

측정하며, 제시된 운영 기간,  $[0, t_j]$ 은 실제 장비가 가동중인 “Power on” 시간만을 의미한다. 그런데 기록시스템의 미비 등으로 인해 실제 확인할 수 있는 고장들은  $s_j$ 시점 이후에 발생된  $n_j - r_j + 1$ 개뿐이며, 그 이전에 발생된  $r_j - 1$ 개의 고장은 확인할 수 없게 된 경우가 많다. 다만, 기록이 정상적으로 시작된 시점의 장비 누적운용시간  $s_j$ 은 알 수 있는 상황이다. 그림 5와 같은 사건이 발생한 자료를 다수의 장비를 대상으로 확장한 것이 본 논문의 분석 대상 자료이며, 형태는 그림 6.으로 표현된다.

그림 6.에서 개별 장비들의 관측 시작시점  $s_j$ 가 모두 상이한 이유는 분석대상 장비들이 동일한 시점에 운영이 시작되지 않기 때문이다. 제조사의 생산능력이나 운용부대의 재편성 등을 고려하여 수년 동안 몇 대씩 점차적으로 배치되는 것이 일반적이다. 또한, 운용기간 동안 장비별로 운용량에 다소 차이가 있을 수 있어 동일시점에 배치된 장비라 하더라도 관측이 시작된 시점  $s_j$ 들은 모두 상이해진다. 그림 7.은 이를 개념적으로

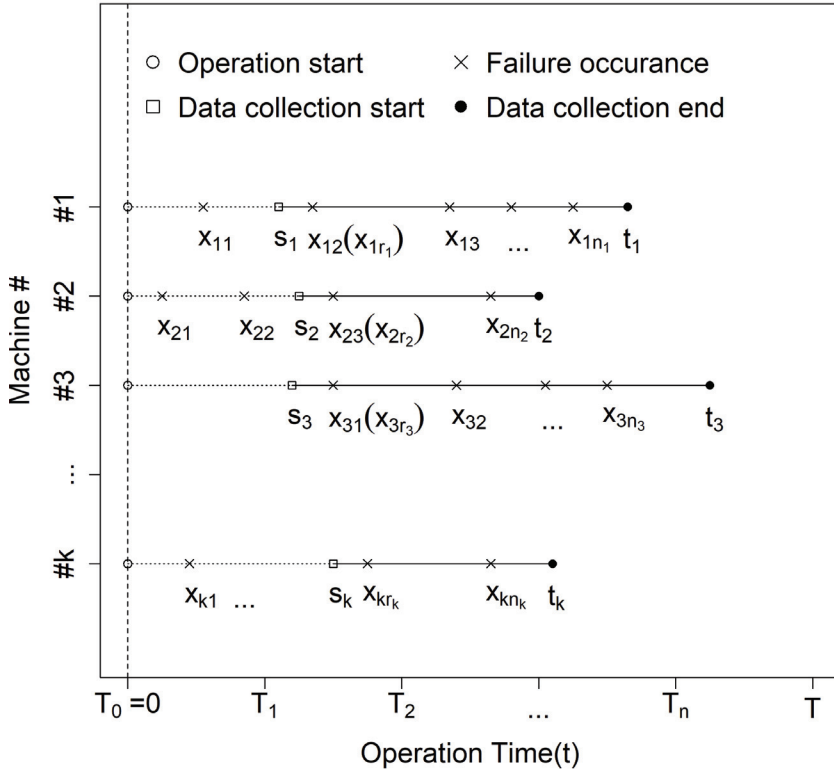


그림 6: 다수 장비의 좌측절단 고장자료 형태

나타낸 것이며, 그림 7.을 누적운용시간 기준으로 재정리 하여 그림 6.을 얻을 수 있는 것이다.

그림 5.의 자료는 표 1.로 재정리 할 수 있으며, 본 논문에서는 표 1. 형태를 분석자료의 기준으로 사용한다.

## 2.2 신뢰도 추세분석의 목적과 영향

신뢰도 추세분석은 단위시간 당 발생하는 고장수의 변화를 분석하는 활동으로 장비 개발과 운용 중에 모두 사용된다. 개발 중에는 신뢰도 목

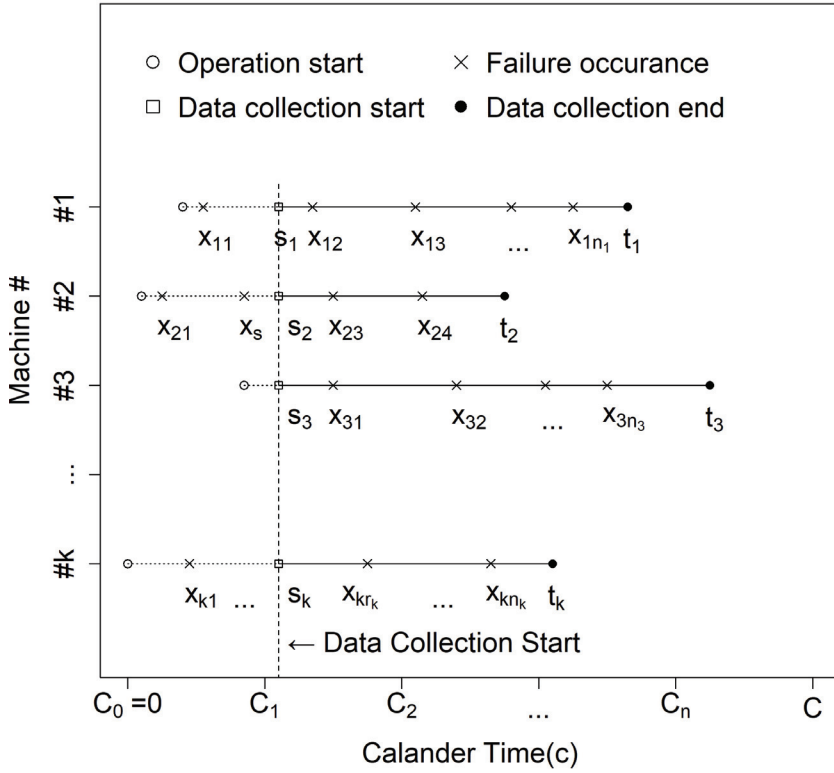


그림 7: 다수 장비의 좌측절단 고장자료 생성개념

표 1: 좌측절단된 다수장비의 고장자료 형태

장비 ID(j)	자료수집 시작 ( $s_j$ )	자료수집 종료 ( $t_j$ )	관측된 고장수 ( $n_j - r_j + 1$ )
1	$s_1$	$t_1$	$n_1 - r_1 + 1$
2	$s_2$	$t_2$	$n_2 - r_2 + 1$
...	...	...	...
N	$s_N$	$t_N$	$n_N - r_N + 1$

표를 달성하기 위한 관리계획을 수립하거나 이를 달성할 수 있는 시점의 추정에 사용된다. 또한, 분석 결과가 관리목표에 부합하지 못하는 경우, 이를 극복하기 위한 자원(예산, 인력, 시간 등) 투입량을 조정하는 근거로도

사용된다.

방산 선진국인 미국에서는 80년대 초반부터 "신뢰성 성장관리"라는 관리기법을 2.3.1절에서 설명될 Power Law 과정을 기반으로 개발하여 사용 중에 있다. 흔히 Duane모형(Duane (1964)), AMSAA 모형(Broemm et al. (2000))으로 알려져 있다. 미군은 위 관리기법을 \$185M 이상의 개발예산이 투입되는 사업에 필수적용을 제도화 하였으며, 시험평가 시 성능과 함께 합·부 판정 요소로 사용한다. 이에 따라 위 신뢰도 추세분석 기법은 EFV(Expenditionary Fighting Vehicle), JSF(Joint Strike Fighter), Osprey 등 최신 대형 무기체계 개발에 적용되고 있다. 그림 8.은 미군의 신뢰도 성장관리 계획 사례 중 하나로 기본적인 계획곡선이 Power Law 형태를 띄고 있음을 알 수 있다. DT(Demonstration Test), IOT(Initial Operation Test), LUT(Limited User Test)등 각종 테스트 시점에서 달성해야 할 신뢰도(MTBF) 수준을 계획하고 있으며, 각 시점에서 확인된 개선사항들을 수정시점을 CAP(Corrective Action Phase)로 표시하는 등 관리를 위한 다양한 요소들을 도식화 하고 있다. 본 논문의 내용과 관계가 약한 사업관리 기법 관련 세부 속성의 설명은 참고문헌인 Handbook (1981); Broemm et al. (2000); Ellner and Herbert (2013) 등으로 대체한다.

장비 운용 중에는 이미 달성된 신뢰도 목표수준이 꾸준히 유지되고 있는지 또는 현재 장비의 신뢰도 수준은 어떠한고 향후에는 어떻게 변화될지, 개선효과의 크기는 얼마나 되는지 등의 확인에 사용된다. 군용장비를 상용품에 비해 긴 기간 동안 사용하는 이유는 상대적으로 고가이고 내구성이 높기 때문이기도 하지만, 지속적인 개선활동을 통해 성능을 유지 또는 향상시키고 있기 때문이다. 또한, 상용품과 달리 엄격하게 계획된 예방정비로 향후 문제가 될 수 있는 부분을 사전에 식별, 조치하기 때문이기도 하다. 이러한 활동의 효과를 확인하는 가장 쉬운 방법으로 신뢰도



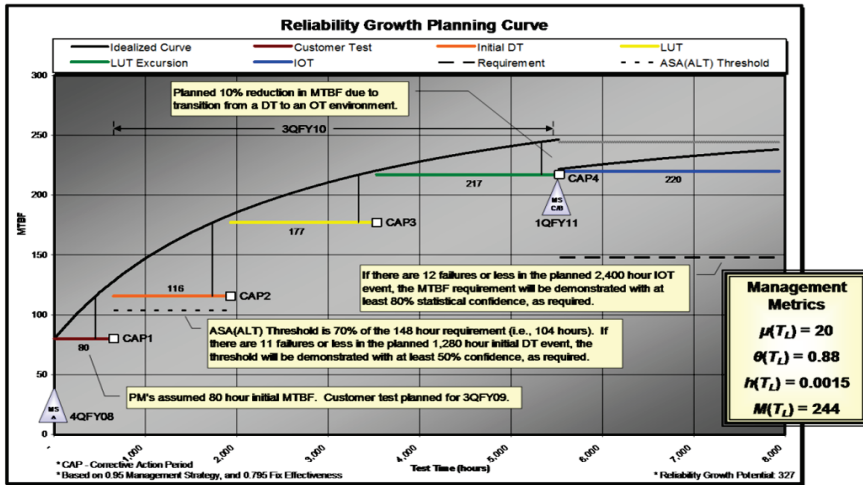


그림 8: 미 육군의 신뢰도 성장관리 계획 사례(Handbook (1981))

추세분석이 사용되어 왔다.

고장간격(Time Between Failures)이 시간에 따라 짧아지고 있지는 않은 지 여부를 통해 현재 수행중인 예방정비의 효과성을 쉽게 확인할 수 있고, 개선활동의 효과성은 고장간격이 시간에 따라 길어지는 지로부터 확인된다. 그림 9.는 이를 개념적으로 표현한 것이다. 특정 장비의 1,800 시간 무렵 수행된 장비개선 즉, 형상변경(Engineering Change Proposal) 결과가 누적 고장수에 어떤 변화를 주었는지를 Power Law 과정으로 분석한 결과이다. 점선으로 표현된 부분이 해당 시간 이전까지의 누적된 자료를 분석한 결과이며, 실선으로 표현된 부분이 이후 누적된 자료까지를 포함한 분석 결과이다. 만약 장비개선이 수행되지 않았다면 1,800시간 이후에도 그림의 점선과 같은 형태로 고장이 발생되어 왔을 것으로 예상되지만, 실제로 발생한 고장은 실선과 같아져 그림의 화살표만큼의 효과를 확인할 수 있게 된다.

예방정비 활동의 효과 측정 역시 유사한 방법으로 가능하다. 만약

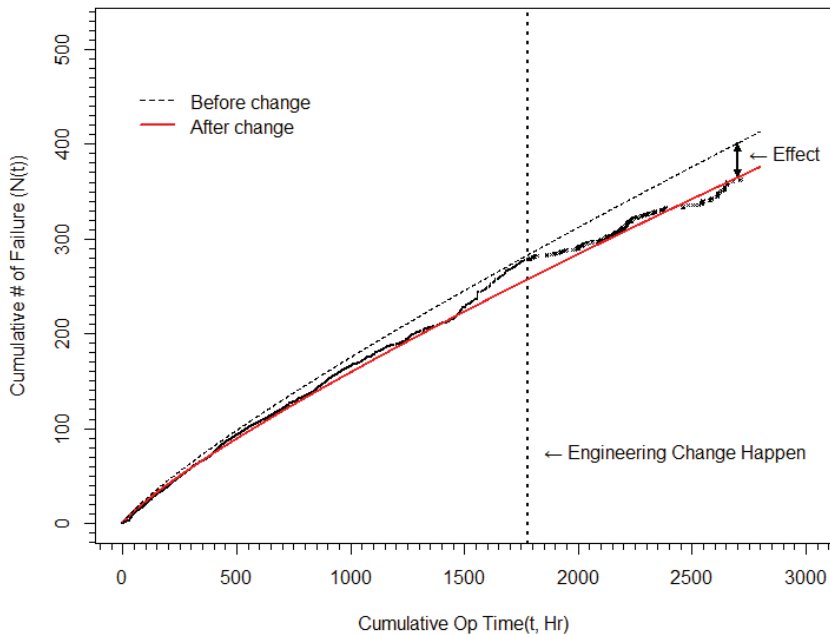


그림 9: 개선효과의 측정 개념

현재 수행중인 예방정비가 효과적이라면, 고장 간 시간 간격은 유지되거나 길어지거나 유지되어야 할 것이다. 한편 이러한 단편적인 효과의 측정 이외에 보다 거시적인 의사결정의 근거자료로 사용되는 경우도 있다. 장비의 폐기나 완전 분해수리, 즉 창정비(Depot Maintenance) 시점 결정이 좋은 사례가 될 수 있다. 군용장비의 경우 대체로 전략적 유효성을 기준으로 장비의 폐기여부를 결정한다. 하지만 전투장비가 아닌 물품 운송차량, 정비 지원 장비 등은 이와 별도로 경제적 효용성을 고려하기도 한다. 경제적 효용성은 결국 정비비용이 향후 일정수준 이상 증가될 것인가를 판단하는 행위이며, 이는 향후 시간에 따른 고장 빈도를 예측하는 행위가 된다.

창정비 시점의 결정도 이와 유사하다. “창정비”는 “장기간 사용으로

열화되어 잦은 고장의 원인이 되는 소모성 자재류(고무패킹, 베어링, 개스킷 등 각종 실링류)를 교환하는 행위”이며 통상 10년 내외의 주기로 수행된다. 완전 분해수리를 해야 하므로, 장기간이 소요되고 이에 따른 비용도 많이 소요된다. 따라서 창정비를 어느 시점에 수행하는 것이 가장 효과적인지를 결정하는 것은 예산확보나 전투력 공백 최소화 측면에서 매우 중요한 문제이다. 따라서 이러한 시점의 결정요소 중 하나로 장비 신뢰성 수준을 고려하고 있다(Brauer and Brauer (1987); Anderson and Neri (2012)). 현 신뢰도 변화 추세를 유지한다면 장비 신뢰도가 목표수준 이하로 열화되는 시점은 언제인지를 확인하여 경제적 효용성과 함께 창정비 시점 판단의 근거로 활용하는 것이다.

국내에서도 전투장비의 신뢰성 제고와 운영중인 장비의 운용성 최적화를 위해 개발 및 운영유지 간 신뢰도 추세분석을 요구하고 있으며, 특히 운영중인 핵심전투장비의 경우 1~2년을 주기로 분석되고 있다. 물론 주요 정비정책의 결정이 요구될 경우 단발성으로 수행되기도 한다.

이 같은 신뢰도 추세분석은 개별장비의 특성을 분석하여 고장시점을 예측하는 상태기반정비(Condition Based Maintenance)와는 거시적인 의사결정을 위한 모형이라는 점에서는 다소 차이를 보인다. 상태기반정비가 개별장비 별 세부적인 고장특성을 분석하여 차별화된 정비정책을 적용하는 것에 목적이 있다면, 본 연구에서 다루고 있는 Power Law 기반의 추세분석은 장비 전체집단의 평균적인 추세를 통해 보다 큰 정비정책을 적용하는 것에 목적이 있다. 장비 전체 집단의 정비비용 추세의 예측이나 창정비 시점의 예측 등에 활용되며, 관련 예산의 반영 등 중장기 계획 수립에 주로 사용이 된다는 점에서 차별화 된다.

이 외에도 신뢰도 추세분석은 여러 분야에서 사용되기 때문에, 확률 모형에서 AI기법까지 다양한 기법들이 사용되어 왔다. 다만, 2.1절에서

설명된 형태의 자료 분석에는 몇 가지 측면에서 사용이 제한되었다. 다음 절에 관련 연구와 시사점들을 간략히 제시한다.

## 2.3 관련 연구

### 2.3.1 연구현황

신뢰도 추세분석은 시스템에서 발생한 고장 시간 간격 또는 단위 시간동안 발생한 고장 개수,  $\lambda(t)$ 의 변동을 분석하는 행위이다. 고장의 발생 원인이나 장비 특성 등을 고려하여 분석대상 자료의 정제가 필요한 경우가 많으나, 정제된 자료는 순수하게 시간에 따른 이벤트 발생 결과에 불과하므로 다양한 통계 모형 적용이 가능하다.

군용장비에 대해 가장 오래되고 널리 사용되는 모형은 고장률의 변화를 식 (2.1)과 같은 강도함수로 표현한 NHPP모형으로, 흔히 Power Law 과정(Power Law Process)로 알려져 있다. Power Law 과정은 Crow(Mitzenmacher (2004))가 제안한 방법으로, 실험식인 Duane(Duane (1964))의 모형을 가장 적합한 확률모형으로 재구성한 모형이다.

$$\lambda(t) = \alpha\beta t^{\beta-1} \quad (2.1)$$

위 모형은 추정이 필요한 모수가 두 개 뿐으로 간단하고, 단순히  $\beta$ 의 크기만으로 분석 대상 자료의 고장률 변동을 표현할 수 있는 장점을 갖는다.  $\beta > 1$ 일 때는 고장이 증가 추세로 장비가 열화되고 있음을,  $\beta < 1$ 일 때는 고장이 감소 상태로 장비가 향상되고 있음을 나타낸다. 특히  $\beta = 1$ 일 때는 고장률이 시간에 따라 변화되지 않아 누적 고장수는 포아송 과정이 되는 특성을 갖는다. 그러나 고장강도함수가 단조형태이기 때문에, 시간

에 따른 고장의 변동을 유연하게 표현하지 못하는 단점이 있고, 고장 발생 원인을 모두 동일하게 가정하는 등 모형의 단순함에 따른 영향이 존재한다. 이 같은 단점에도 불구하고 모형의 단순성으로 인해 신뢰도 추세분석에는 Power Law 모형이 꾸준히 사용되어 왔으며, 일부 단점을 보완하고 실효성을 높이기 위한 연구 활동이 다수 확인된다. 각 연구들은 일정한 방향성을 갖고 있으며, 과거 연구 활동의 방향성에 따라 다음과 같이 요약될 수 있다.

우선 고장의 발생이나 회복에 확률모형을 사용하였으나 Power Law와 Intensity의 형태가 다르거나 가정이나 관점이 다른 연구들이 확인된다. NHPP 모형의 경우  $T_i$ 시점까지의 누적 고장수  $\Lambda(T_i)$ 의 관계인 식 (2.2)의 관계로부터, 지수분포 형태를 띄는 Shock이 매번 고장이 발생됨을 모사한다.

$$\Lambda(T_i) - \Lambda(T_{i-1}) \sim \text{Exp}(1) \quad (2.2)$$

그러나 Shock이 있을 때마다 고장이 반드시 발생하는 것은 아니기 때문에, Lakey (1992), Muralidharan (2002), Black and Rigdon (1996) 등은 고장을 유발하는 스트레스의 중첩을 고려하는 MPLP(Modulated Power Law Process)를 제안하였다. MPLP에서  $\Lambda(T_i)$ 의 관계는 식 (2.3)으로 표현된다.  $\Gamma(\kappa, 1)$  분포는  $\text{Exp}(1)$ 이  $\kappa$ 개 중첩된 분포이므로, 이는  $\kappa$ 회의 지수분포 형태 Shock이 누적된 후에야 비로소 고장이 발생됨을 의미한다. 실제로 하나의 고장이 발생은 지속적인 스트레스 누적에 의한 현상이 대부분이므로, 기계류 구성품의 Shock과 고장발생을 모사하기에 유리하다.

$$\Lambda(T_i) - \Lambda(T_{i-1}) \sim \Gamma(\kappa, 1) \quad (2.3)$$

또 다른 형태의 모형으로, 고장 시간이 아닌 고장 간 시간  $X_j$ 를 이용하여 고장의 발생이 아닌 수리효과 측면의 모형도 존재한다. Arab et al. (2012)은 고장 간 시간  $X_j$ 의 관계를 식 (2.4)로 표현하고, 이를 PEXP(Piecewise EXPonential)라 하였다. 위 모형은  $X_j$ 에 관한 이산모형으로, 각  $X_j$ 들은 i.i.d로 가정한다. 기본적으로 위 모형은 Power Law 과정과 유사하나, Power Law가 고장 추세의 변화를 시간 또는 제품의 열화에서 찾은 모형인 반면, MEXP는 고장률의 변동을 실제 고장이 발생된 시간에서 찾은 모형이라 할 수 있다. Power Law와 유사하게 식 (2.4)에서  $\delta$ 의 크기에 의해 고장 발생 후 수리(or 개선)효과의 유무를 판단한다. 즉,  $\delta > 1$ 일 때는  $X_j$ 가 늘어나므로, 수리 또는 개선의 효과가 반영되어 신뢰도 향상을,  $\delta < 1$ 일 때는 개선 효과가 이전보다 떨어짐을,  $\delta = 1$ 일 때는 개선효과가 없음을 의미하게 된다.

$$E(X_j) = \frac{\delta}{\mu} j^{\delta-1} \quad (2.4)$$

그림 10.은 PEXP의 개념 설명을 위해 Arab et al. (2012)이 제시한 그림으로,  $\delta > 1$ 일 때, 고장강도(Failure intensity) 변화를 나타내며, 모형의 특성인 고장 발생 시점마다 이산적으로 변화하는 고장률이 확인된다.

NHPP 모형이기는 하나, Power Law와 다른 형태의 Intensity를 사용하거나, 몇 개의 강도함수를 중첩하여 사용하기도 하였다. Ascher and Kobbacy (1995)는 Cox and Lewis (1966)가 제안한 식 (2.5)와 같은 Log-Linear 형태의 강도함수를 사용하여 시간에 따른 고장수 변화를 모형화 하였고, Pulcini (2001)와 Mun et al. (2013)는 각각 Power Law와 Log-Linear 형태의 강도함수를 시간대별로 중첩하여 특정시점 이후의 고장 강도함수 변화를 반영하고자 하였다.

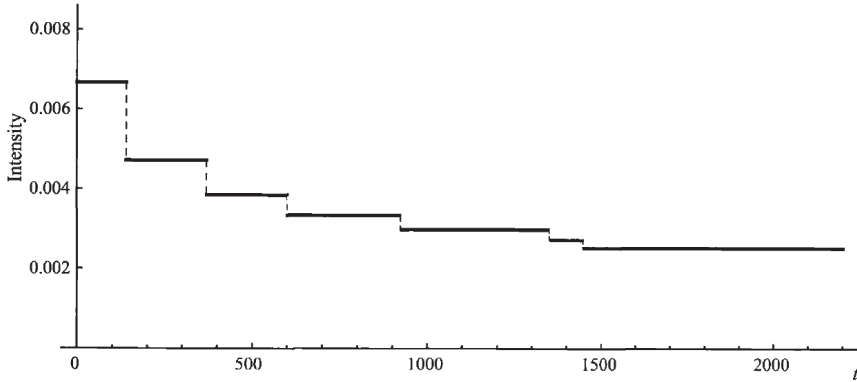


그림 10: PEXP Model 개념, ( $\delta > 1$ 일 때, Arab et al. (2012))

$$\lambda(t) = \text{Exp}(\gamma_0 + \gamma_1 t) \quad (2.5)$$

이 외에도, Calabria and Pulcini (1997), Hamada et al. (2008) 등은 각 모형들에 Bayesian 개념 적용을 통해 모형의 정교화와 효용성 향상을 시도하였다.

한편 Power Law 과정의 기본특성이 아닌 이를 다수 장비로 확장하기 위한 노력도 확인된다. 일반적으로 원자력발전소, 플랜트 등 Power Law 과정은 단일장비를 대상으로 하나(Sun et al. (2005)), 다수 장비에서 확보된 고장자료를 활용할 때는 각 장비에서 발생하는 고장은 서로 독립이고 각 장비들의 특성은 동일함을 가정하는 것이 일반적이다. 만약 모든 장비의 강도함수,  $\lambda(t)$ 가 식 (2.1)과 같고 정시중단(Type I Censoring) 상황에서 총  $N$ 개 장비로부터  $j$ 번째 장비의  $i$ 번째 고장인  $x_{ji}$ 를 장비별로  $n_j$ 개씩 확보하였을 때, Likelihood는 식 (2.6)으로 표현된다. 실제  $\lambda(t)$ 의 모수는 식 (2.6)에 score 함수 ( $\frac{\partial \ln L}{\partial \alpha} = 0, \frac{\partial \ln L}{\partial \beta} = 0$ )로 부터 손쉽게 얻을 수 있다.

$$L = \prod_{j=1}^N \left( \left( \prod_{i=1}^{n_j} \lambda(x_{ji}) \right) \text{Exp} \left[ - \int_0^t \lambda(u) du \right] \right) \quad (2.6)$$

그러나 모든 장비가 항상 동일한 강도함수를 가지지 않을 수도 있다. 동일한 장비라 하더라도 사용자나 사용 환경에 따라 고장의 추세가 다를 수 있다. 이러한 장비의 동질성 가정을 저해하는 요소는 사용빈도, 노출온도, 사용 군 등 기록이나 식별이 가능한 요소가 있는 반면, 연도별 제작기술 변화, 사용자 습관 등 상이한 특성을 나타낼 것으로 예상되나 기록으로 구분하기 어려운 요소도 있다. 전자는 공변량(Covariate)이라는 요소로 구분하여 모형에서 고려하는 것이 일반적이며, 후자는 Frailty라는 용어로 정의되고 모형에 반영되어 왔다.

전자의 경우, 주로 Power Law 과정과 같은 NHPP 모형보다는 Cox (1992)가 제안한 PHM(Proportional Hazard Model) 계열 모형들이 많다. Kumar (1995), You and Meng (2012)는 PHM을 이용하여 수리가능한 시스템을 분석한 사례들이라 할 수 있다.

후자의 경우, 기본적인 Power Law 강도함수에 Frailty를 의미하는 변수를 추가로 고려하는 방식으로 수행된다. Lindqvist (2006)의 연구에 의하면, 실제로 Frailty가 존재하나 존재를 고려하지 않은 결과가 다수 존재하는 것으로 알려졌으며, 이를 고려하지 않음에 따라 결과가 최적 값이 아니거나 완전히 틀린 결론에 이르기도 함을 지적한 바 있다. 이는 식 (2.7)과 같이 수정된 강도함수를 사용함으로써 해소될 수 있다.

$$\lambda(t) = Z\alpha\beta t^{\beta-1} \quad (2.7)$$



물론,  $Z$ 를 고려함에 있어 분포함수를 가정하는 방법(Aalen et al. (2008); Lawless (1987)), 특히 사용의 편리함 때문에 감마분포를 사용하는 경우가 많으나(Asfaw and Lindqvist (2015)), 최근 Slimacek and Lindqvist (2016)는 비모수적 방법의 사용을 시도하기도 하였다. 이는 공변량을 사용하는 경우와 달리 Frailty는 실제  $Z$  자체의 추정치는 무의미하고  $Z$ 의 크기, 즉, 분산(Variance)만을 중요시하기 때문이다.

이 외에도 확률모형 이외의 방법을 이용하여 신뢰도 추이를 분석하려는 시도는 매우 다양하다. Ho and Xie (1998)와 Ho et al. (2002)은 일정기간 동안에 발생한 고장을 시계열 형태의 고장으로 재가공하여 Box-Jenkins 모형으로 분석을 시도한 바 있고, Hong and Pai (2006), Fuqing et al. (2010)는 지지벡터머신(Support Vector Machine)을 사용하여 신뢰도 추세를 분석하려 하였다. 또한 Gat (2014)은 Yule Process를 이용하여 좌측절단 상태로 기록된 상수도관의 고장 추세 분석을 시도하기도 하였다.

이처럼 다양한 모형의 존재에도 불구하고, 사용의 편의성과 기존 결과와의 호환성에 장점이 있기 때문에, Power Law 과정은 자료의 특성과 목적에 따라 불완전 수리(Imperfect repair)를 고려하거나(Dijoux et al. (2016)), 부족한 자료의 보완을 위해 Bootstrapping을 사용하는 등(Gilardoni et al. (2013)) 약간씩의 변형을 주고 있지만, 여전히 신뢰도 추세분석에 사용되고 있다. 이에 본 연구에서도 Power Law 과정을 기본모형으로 사용하고자 하였다.

### 2.3.2 기존연구의 한계점

실제 장비 사용 중에 수집된 자료들에 위 모형을 적용하는 것은 모형의 개선을 하는 것 이상으로 어려운 문제이다. 우선 신뢰도 추세분석을 하여 정비정책을 결정할 만큼의 중요성을 가지는 제품들은 대부분 대량

생산되는 제품이기보다는 대부분 고가의 플랜트나 단독시스템인 경우가 많다. 이 때문에, 위에서 언급된 대부분의 모형이 장비군 보다는 단독장비나 소프트웨어를 대상으로 한다. 개발단계에서 사용되는 몇 개의 시제품 자료를 이용하기 위해, Crow (1975)가 MLE 기반의 모형을 제시한 바 있으나, 운용단계의 대량으로 사용 중인 장비 군에서 발생하는 자료를 분석할 수 있는 방법은 특별히 제안되어있지 않다. 즉, 고가의 장비이면서 어느 정도의 개체수량이 존재하는 전차, 장갑차와 같은 무기체계에 적용할 수 있는 방법은 존재하지 않는다. 또한, 위 모형들은 대상물의 운영시작 시점부터 수집된 정상자료를 사용함을 전제한다.

단일장비에서 발생하는 부분적 자료누락인 경우 Crow and Basu (1988)의 방법과 같이 자료 유실 전후의 고장추세를 내삽 하는 방법의 사용이 가능하다. 하지만 장비군 전체에서 대량으로 자료누락이 발생하거나, 특히 수집된 기간 중 중간 중간의 일부자료가 누락된 것이 아닌, 일정시간 이전 자료 전체가 누락되는 경우, 즉 좌측절단 상황은 처리가 더 복잡하다.

### 2.3.3 시사점과 연구방안 수립

선행 연구결과의 확인 결과, 본 연구에서 효과적으로 활용 가능한 연구로 Yu et al. (2008); Slimacek and Lindqvist (2016, 2017)을 확인할 수 있었다. 위 연구들에서 사용된 방법과 개념을 활용하고 각 연구의 제한사항을 보완할 경우 본 연구의 목표를 보다 빠르게 달성할 수 있다고 판단하게 되었다.

현재로서 가장 유사한 연구는 Yu et al. (2008)이었다. 위 연구는 좌측절단이 발생된 단일 장비의 자료를 Power Law 과정으로 분석할 수 있는 방법을 제안하고 있다. 제안된 결과는 좌측절단 상황을 고려한 거의 유일한 연구이나, 두 가지 측면에서 본 연구에 직접 적용은 곤란하였다. 우선

손실된 자료의 개수를 확인할 수 있어야 한다는 점이다. 즉, 그림 5.에서 최초 관측된 고장  $x_{jr_j}$ 가 전체  $n_j$ 개의 고장 중  $r_j$ 번째 고장임을 알 수 있어야 한다. 그러나 현실적으로 관측이 시작된 시점  $s_j$ 를 알 수는 있으나,  $s_j$  이전에 몇 개의 고장이 손실되었는 지 알 가능성은 거의 없다. 따라서 수학적 모형으로는 의미를 가질 수 있으나, 실무적 활용성은 낮다고 할 수 있다. 또한, 단일 장비를 대상으로 한 연구로 본 논문에서 다루고자 하는 그림 7.과 같은 형태의 다수 장비에서 발생한 고장자료를 분석하려면 우도함수와 모수 추정식의 재산출이 요구되었다.

본 연구에서는 위와 같은 제한점을 해결하기 위해 ML 추정법을 사용하여 우도함수와 모수 추정식을 재 산출 하였으며, 관측이 시작된 고장의 순서(order),  $r_j$ 를 추정하기 위해 E-M 알고리즘을 사용키로 하였다. 이 같은 접근법은 기존 연구결과를 충분히 활용하여 논리적 타당성을 얻을 수 있음과 동시에 실용성도 갖출 수 있을 것으로 판단하였다. 그러나 근본적으로 다수 장비의 이질성(Heterogeneity)은 고려할 수 없는 단점의 극복 필요성도 인식하게 되었다.

다음으로 활용 가능한 연구에는 Slimacek and Lindqvist (2016, 2017)를 확인할 수 있었다. 위 연구들은 다수장비의 이질성을 고려하기 위해, 관측 가능성 유·무에 따라 각각 Frailty와 공변량(Covariates)이라는 요소를 고려하고 있다. 기본적으로 본 연구에서 요구하는 다수장비를 대상으로 하고 있고, 무엇보다, 앞서 언급된 Yu et al. (2008)의 연구와 동일한 Power Law 형태의 강도함수를 갖는 NHPP를 기본모형으로 하고 있다는 점도 확인할 수 있었다. 다만, 좌측절단 상황이 모형에 반영되지 않았다는 점만이 본 연구와 약간 차이가 있는 연구라 할 수 있다. 이에 따라, 만약 다수 장비의 동질성을 전제로 구축한 모형이 있다면 강도함수의 변화를 통해 손쉽게 위 모형의 결과를 활용할 수 있다는 판단을 하게 되었다.

이상의 시사점들로부터, 우선 Yu et al. (2008)의 연구를 이용, 다수장비의 동질성을 전제로 한 “기본모형”을 구축 후 강도함수의 변형을 통해 “확장모형”으로 일반화 하는 방법을 연구 전략으로 사용키로 하였다.

## 제 3 장

### 신뢰도 추세분석 기본모형

Yu et al. (2008)의 모형은 단점은 최초로 관측된 고장의 순서(order), 즉,  $r_j$ 가 몇 번째 고장인지를 알아야 한다는 사실이다. 하지만 현실에서는 이를 알 가능성은 거의 없다. 다행히도 표 1.에서 알 수 있듯이 본 논문의 대상 자료는 관측 시작시점  $s_j$ 들을 확인 할 수 있다. 따라서 위 정보를 활용, E-M 알고리즘으로  $j$ 번째 각 장비들의 관측되지 않은 시점의 고장수  $\hat{r}_j$ 를 추정하여 Yu et al. (2008)의 제한점을 극복할 수 있는 방안을 제안한다 (Na and Chang (2017)).

E-M 알고리즘(Little and Rubin (2014))은 손실된 자료를 추정된 기대값으로 대체하는 반복적 과정으로 식 (3.1)과 같은 E-Step과 M-Step 두 부분으로 구성된다.

$$\begin{aligned} E - step & : Q(\theta|\theta^{(k)}) = E_{Z|X, \theta^{(k)}} [L(\theta|X, Z)] \\ M - step : & : \theta^{(k+1)} = \underset{\theta}{\operatorname{argmax}} Q(\theta|\theta^{(k)}) \end{aligned} \quad (3.1)$$

E-step (Expectation step)에서는 관측된 자료  $X$ 와 현재까지 추정된 최신 모수  $\theta^{(k)}$ 를 이용하여 손실된 자료  $Z$ 의 기대값을 구하는 과정이며, M-step (Maximization step)은 손실된 자료를 기대값으로 채운 자료를 이용하여 우도함수(Likelihood function)  $L(\theta|X, Z)$ 를 최대화 하는 모수  $\theta^{(k+1)}$ 을 산출하는 과정이다. 이 과정은 모수추정치가 수렴될 때( $\theta^{(k+1)} - \theta^{(k)} < \epsilon$ )까 지 반복적으로 수행된다. 위 방법은 자료의 손실을 보상하는 손쉬운 방법

으로, 손실이 발생된 자료를 이용한 분석에 폭넓게 사용된다. 다만, 전역 최적화 값을 보장하지는 않는 단점이 있다. 따라서 본 논문에서는 이를 보완하기 위해 초기값으로 임의의 값을 사용하는 방식이 아닌 관측된 자료의 형태로부터 일정수준 최적 값에 근사된 초기값을 산출하여 E-M 알고리즘의 초기값으로 사용하여 이를 일정 수준 보장하고자 한다. 이를 통해 산출된 해는 반드시 최적 값은 아닐 수 있으나 최적 값에 매우 근사한 값임을 보장할 수 있을 것으로 판단하였다. E-M 알고리즘 적용과정 중 실제 수치해석에 의한 모수추정에는 Newton-Raphson 방법을 적용하였다.

기본모형의 모수추정 절차는 그림 11.과 같으며, 각 부분에 대한 세부 사항은 이어지는 장절에서 설명한다.

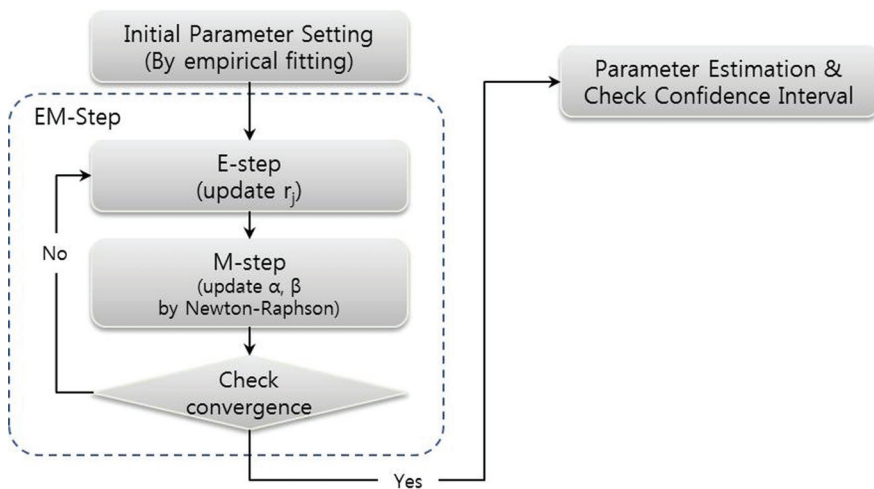


그림 11: 기본모형의 모수추정 절차

### 3.1 좌측절단을 고려한 모수추정 방안

그림 5에서,  $j$ 번째 장비의 고장자료  $X_j = (x_{j,r_j}, \dots, x_{j,n_j})$ 는 정시중단 (Type I Censored) 자료로, 실제 고장의 기록은  $s_j$ 시점부터  $t_j$ 시점까지이다. 이 중  $s_j$ 까지 총  $(r_j - 1)$ 개의 고장은 기록이 남아있지 않으며,  $s_j$ 이후 발생된  $(n_j - r_j + 1)$ 개의 고장만이 기록된 상태이다. 만약 유실된 고장의 개수를 알 수 있다면, Little and Rubin (2014)의 아이디어를 일부 사용할 수 있다.  $N[t_0, t_1]$ 을  $[t_0, t_1]$ 구간에서 발생된 고장수로 정의할 때, Yu et al. (2008)는 단일장비에서  $[0, t]$ 동안 발생한 고장  $(x_1, \dots, x_n)$ 의 확률밀도함수 (pdf)를 식 (3.2)로 표현하였다.

$$f(x_1, \dots, x_n; N[0, t] = n) = (\alpha\beta)^n \exp(-\alpha t^\beta) \prod_{i=1}^n x_i^{\beta-1} \quad (3.2)$$

이와 유사하게,  $s_j$  이전에 발생된  $(r_j - 1)$ 개의 고장을 인지할 수 있음을 전제하면  $j$ 번째 장비의 확률밀도함수는 다음과 같은 일련의 과정으로 산출된다.

NHPP를 따르는  $N[t_0, t_1]$ 는  $k = 0, 1, \dots$ 일 때, 아래와 같은 확률질량함수 (Probability mass function, pmf)를 만족한다. ( $\int_{t_0}^{t_1} \lambda(y) dy = \Lambda[t_0, t_1]$ )

$$P(N[t_0, t_1] = k) = \frac{\Lambda[t_0, t_1]^k}{\Gamma(k+1)} \exp(-\Lambda[t_0, t_1]) \quad (3.3)$$

만약,  $\tau_x$ 가  $x$ 이후 고장 발생 이전까지의 시간이라면,

$$P(\tau_x > t) = P(N[x, x+t] = 0) = \exp(-\Lambda[x, x+t]). \quad (3.4)$$

를 만족하고, 이로부터  $\tau_x$ 의 확률밀도함수 (Probability density function,

pdf)는 아래와 같이 표현된다.  $\Gamma(r_j) = (r_j - 1)!$ 를 의미한다.

$$f_{\tau_x}(t) = \lambda(x+t) \exp(-\Lambda[x, x+t]) \quad (3.5)$$

따라서, 식 (2.1), (3.3), (3.4), and (3.5)로부터  $N[s_j, t_j]$ 의 확률밀도함수는 식 (3.6)으로 표현될 수 있다.

$$\begin{aligned} f(X_j; s_j, t_j) &= f(x_{j,r_j}, \dots, x_{j,n_j}; N[0, s_j] = r_j - 1, N[s_j, t_j] = n_j - r_j + 1) \\ &= P(N[0, s_j] = r_j - 1) \cdot f_{\tau_{s_j}}(x_{j,r_j} - s_j) \\ &\quad \times \prod_{k=0}^{n_j - r_j - 1} f_{\tau_{x_{j,r_j+k}}}(x_{j,r_j+k+1} - x_{j,r_j+k}) \cdot P(\tau_{x_{j,n_j}} > t_j - x_{j,n_j}) \\ &= \frac{\Lambda[0, s_j]^{r_j-1}}{\Gamma(r_j)} \exp(-\Lambda[0, s_j]) \cdot \lambda(x_{j,r_j}) \exp(-\Lambda[s_j, x_{j,r_j}]) \\ &\quad \times \left( \prod_{k=0}^{n_j - r_j - 1} \lambda(x_{j,r_j+k+1}) \right) \exp \left( - \sum_{k=0}^{n_j - r_j - 1} \Lambda[x_{j,r_j+k}, x_{j,r_j+k+1}] \right) \\ &\quad \times \exp(-\Lambda[x_{j,n_j}, t_j]) \\ &= \frac{\Lambda[0, s_j]^{r_j-1}}{\Gamma(r_j)} \exp(-\Lambda[0, t_j]) \prod_{k=r_j}^{n_j} \lambda(x_{j,k}) \\ &= \frac{1}{\Gamma(r_j)} \alpha^{n_j} \beta^{n_j - r_j + 1} s_j^{\beta(r_j-1)} \exp(-\alpha s_j^\beta) \prod_{k=r_j}^{n_j} x_{j,k}^{\beta-1} \end{aligned} \quad (3.6)$$

식 (3.6)을 이용하여  $N$ 개 장비 모두를 고려한 로그우도는 식 (3.7)과 같고, Power Law 과정의 모수  $\alpha, \beta$ 의 ML 추정치는 아래 식의 score 함수인



식 (3.8)을 0으로 하는  $\alpha, \beta$ 에서 얻을 수 있다.

$$\begin{aligned}
\ln L(\alpha, \beta; X_1, \dots, X_N) &= \sum_{j=1}^N \ln f(X_j; s_j, t_j) \\
&= \ln \alpha \sum_{j=1}^N n_j - \alpha \sum_{j=1}^N t_j^\beta + \ln \beta \sum_{j=1}^N (n_j - r_j + 1) \\
&\quad + \beta \sum_{j=1}^N \left( (r_j - 1) \ln s_j + \sum_{k=r_j}^{n_j} \ln x_{j,k} \right) \\
&\quad - \sum_{j=1}^N \left( \sum_{k=r_j}^{n_j} \ln x_{j,k} + \Gamma(r_j) \right) \tag{3.7}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \alpha} &= \frac{1}{\alpha} \sum_{j=1}^N n_j - \sum_{j=1}^N t_j^\beta \\
\frac{\partial \ln L}{\partial \beta} &= -\alpha \sum_{j=1}^N t_j^\beta \ln t_j + \frac{1}{\beta} \sum_{j=1}^N (n_j - r_j + 1) \\
&\quad + \sum_{j=1}^N \left( (r_j - 1) \ln s_j + \sum_{k=r_j}^{n_j} \ln x_{j,k} \right) \tag{3.8}
\end{aligned}$$

이상의 과정을 통해 얻은  $\alpha, \beta$ 의 추정값  $\hat{\alpha}, \hat{\beta}$ 는 각각 식 (3.9), 식 (3.10)과

같다.

$$\hat{\alpha} = \frac{\sum_{j=1}^N n_j}{\sum_{j=1}^N t_j^{\hat{\beta}}} \quad (3.9)$$

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{j=1}^N n_j - \sum_{j=1}^N r_j + N}{\hat{\alpha} \sum_{j=1}^N t_j^{\hat{\beta}} \ln t_j - \sum_{j=1}^N (r_j - 1) \ln s_j + \sum_{j=1}^N \sum_{k=r_j}^{n_j} x_{j,k}} \\ &= \frac{\sum_{j=1}^N n_j - \sum_{j=1}^N r_j + N}{\sum_{j=1}^N n_j \sum_{j=1}^N \omega_j \ln t_j - \sum_{j=1}^N (r_j - 1) \ln s_j + \sum_{j=1}^N \sum_{k=r_j}^{n_j} x_{j,k}} \\ s.t. \quad &\omega_j = t_j^{\hat{\beta}} / \sum_{j=1}^N t_j^{\hat{\beta}}, \quad \sum_{j=1}^N \omega_j = 1 \end{aligned} \quad (3.10)$$

위 식들은  $\hat{\beta}$ 에 대한 식들이나, 식 (3.10)은 closed-form으로 정리되지 않기 때문에, 수치 해석적 방법의 사용이 요구되며, 수치해석 방법으로는 Newton-Raphson(Süli and Mayers (2003))을 적용하였다. 이후 추정된  $\hat{\beta}$ 를 식 (3.9)에 적용하여  $\hat{\alpha}$ 를 얻을 수 있다. Newton-Raphson 적용과정은 다음과 같다.

Newton-Raphson은 MLE의 성질을 이용하는 수치해석 방법으로, MLE가 단조함수 형태이고 연속함수인 성질을 이용한다. 본 논문에 사용된 우도함수는 2개의 모수를 가지므로 2차원 Newton-Raphson을 적용한다.  $\theta = (\alpha, \beta)'$ ,  $F(\theta) = \left( \frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta} \right)$  일때,  $F(\theta) = 0$ 을 만족하는  $\theta$ 는 Jacobian,  $J(\theta)$ 를 이용한 식 (3.11)로 구할 수 있다.

$$\theta_{k+1} = \theta_k - J(\theta_k)^{-1} F(\theta_k) \quad (3.11)$$

이 때, Power Law 과정의 모수  $\alpha, \beta > 0$  이므로,

$$\theta' = (\ln \alpha, \ln \beta) = (a, b)$$

인 변환을 이용한다. 변환 후에는

$$\begin{aligned}\frac{\partial \alpha}{\partial a} &= \frac{1}{\alpha} \\ \frac{\partial \beta}{\partial b} &= \frac{1}{\beta}\end{aligned}$$

가 모두 양수이므로, 식 (3.12)와 같이 chain-rule에 의해,  $F(\theta) = 0$ 은  $G(\theta') = 0$ 과 동일한 문제가 된다. 따라서 실제로는  $G(\theta') = 0$ 의 해를 찾고 이를 역 변환 하여  $\alpha, \beta$ 를 얻게 된다.

$$\begin{aligned}F(\theta) &= \left( \frac{\partial L}{\partial a} \frac{\partial a}{\partial \alpha}, \frac{\partial L}{\partial b} \frac{\partial b}{\partial \beta} \right) = 0 \\ &\Leftrightarrow \\ G(\theta') &= \left( \frac{\partial L}{\partial a}, \frac{\partial L}{\partial b} \right) = 0\end{aligned}\tag{3.12}$$

$G(\theta)$ 와  $J(\theta)$ 의 세부 산식은 각 모수를 로그 변환한

$$a = \ln \alpha, b = \ln \beta$$

와

$$\begin{aligned}K &= \sum_{j=1}^N (n_j - r_j + 1) \\ G(\theta) &= \begin{bmatrix} A \\ B \end{bmatrix} \\ J(\theta) &= \begin{bmatrix} C & D \\ D & E \end{bmatrix}\end{aligned}\tag{3.13}$$

에 대해 식 (3.14)에서 식 (3.17)과 같이 나타내어 진다.

$$A : \frac{\partial L}{\partial a} = \sum_{j=1}^N n_j - \alpha \sum_{j=1}^N t_j^\beta \quad (3.14)$$

$$B : \frac{\partial L}{\partial b} = K - \alpha \beta \sum_{j=1}^N t_j^\beta \ln t_j + \beta \sum_{j=1}^N \left[ (r_j - 1) \ln s_j + \sum_{k=r_j}^{n_j} \ln x_{j,k} \right] \quad (3.15)$$

$$C : \frac{\partial^2 L}{\partial a^2} = -\alpha \sum_{j=1}^N t_j^\beta \quad (3.16)$$

$$D : \frac{\partial^2 L}{\partial a \partial b} = -\alpha \beta \sum_{j=1}^N t_j^\beta \ln t_j$$

$$E : \frac{\partial^2 L}{\partial b^2} = -\alpha \beta \sum_{j=1}^N t_j^\beta \ln t_j - \alpha \beta^2 \sum_{j=1}^N t_j^\beta (\ln t_j)^2 + \beta \sum_{j=1}^N \left[ (r_j - 1) \ln s_j + \sum_{k=r_j}^{n_j} \ln x_{j,k} \right] \quad (3.17)$$

위 식들을 통해 추정된 모수의 확보를 전제로, Power Law 과정의 기본성질을 통해 신뢰도 관련 인자들을 추가로 산출할 수 있다. 우선 각 시점에서의 고장률은 Power Law 과정의 강도함수이므로, 식 (3.18)과 같다. 또한, 각 시점까지의 누적 고장수는 NHPP의 기본성질에 따라 식 (3.19)로 표현된다. 마지막으로 각 시점에서 확인되는 고장 간 평균 시간은 식 (3.20)으로 표현된다.

$$\hat{\lambda}(t) = \hat{\alpha} \hat{\beta} t^{\hat{\beta}-1} \quad (3.18)$$

$$\hat{m}(t) = \hat{\alpha} t^{\hat{\beta}} \quad (3.19)$$

$$\widehat{\text{MTBF}}(t) = \frac{1}{\hat{\alpha} \hat{\beta} t^{\hat{\beta}-1}} \quad (3.20)$$

일반적인 신뢰성공학(O'Connor et al. (2012); Ramakumar (1993))에서는 식 (3.18)과 같이 정의된 고장률을 순간고장률(Instantaneous failure rate,  $\lambda_i(t)$ )로, 누적 고장수를 활용한 식 (3.21)을 누적고장률(Cumulative failure rate,  $\lambda_c(t)$ )로 정의하기도 하나, 통상 고장률은  $t$ 시점에서 경험할 수 있는 순간고장률을 의미하므로, 본 논문에서는 식 (3.18)만을 고려한다.

$$\lambda_c(t) = \frac{\hat{m}(t)}{t} = \frac{\hat{\alpha} t^{\hat{\beta}}}{t} \quad (3.21)$$

## 3.2 모수추정

3.1절에서 확보된 식을 이용하기 위해서는  $N$ 개의 장비에 대해  $r_j$ 를 알고 있어야 한다. 앞서 언급한 바와 같이  $\hat{r}_j$ 는 E-M 알고리즘을 이용하여 산출한다. 이 과정에서 E-M 알고리즘의 단점인 전역최적화를 보장하지 못함을 보상하기 위해 E-M 알고리즘의 초기값으로는 경험적(Empirical) 방법으로 계산된 값을 사용한다. 이 외에도 양질의 초기값을 산출하기 위한 유사한 방법으로 좌측절단 자료를 좌측절단이 발생되지 않은 자료로 간주한 후 Power Law 모형으로 분석한 값을 적용하는 방법도 고려할 수 있다. 이는  $s_j$ 시점을 장비가 운영을 시작한 시점인  $t = 0$ 으로 간주하는 것과 같다. 만약, 장비 전체적으로 좌측절단이 발생한 시점인  $s_j$ 가 유사하다면 본 연구와 유사한 품질의 초기값을 산출할 것이다. 그러나 경우에 따라  $s_j$ 의 차이가 큰 경우도 있을 수 있다. 이 경우 고장 추세의 왜곡이 발생할

가능성이 있으므로, 추가 정보인  $s_j$ 를 활용할 수 있는 3.2.1을 초기값 산출 방법으로 사용하였다.

### 3.2.1 추정의 실효성을 높일 수 있는 초기값 추정

모든 장비 중에 최초관측시간이 가장 짧은 시점을  $s = \min s_j$ , 관측중단 시간이 가장 늦은 시점을  $T = \max t_j$ 라 하고,  $[s, T]$ 를  $\{[T_{m-1}, T_m] : m = 1, \dots, L, T_0 = s, T_L = T\}$ 과 같이  $L$ 개의 소구간으로 분할한다고 가정한다. 이 때,  $Y_m$ 을  $[T_m, T_{m+1}]$ 에서 운용중인 장비에서 관측된 고장 개수라 하면,  $Y_m$ 은

$$Y_m = \sum_{j \in M_m} \sum_{k=1}^{n^j} \mathbf{1}\{x_{jk} \in [T_m, T_{m+1}]\} \quad (3.22)$$

로 표현 될 수 있다.(단,  $M_m \subset \{1, \dots, N\}$ 은 구간  $[T_m, T_{m+1}]$ 에서 작동 중인 장비) 만약  $|M_m|$ 이 충분히 크다면, 대수의 법칙(the law of large numbers)에 따라,

$$Y_m/|M_m| \rightarrow E[N[T_m, T_{m+1}]] \quad (3.23)$$

으로 근사할 수 있다. 또한,  $\Delta T = T_{m+1} - T_m$ 라면,

$$\frac{Y_m}{|M_{m+1}|} \approx E[N[T_m, T_{m+1}]] = \int_{T_m}^{T_{m+1}} \lambda(s) ds \approx \lambda(T_m) \Delta T \quad (3.24)$$

와 같은 관계가 성립한다. 이를 그림으로 나타내면 그림 12.와 같다. 그림 12.에서 사각형 음영부분은  $\lambda(T_m) \Delta T$ 을 나타내며, 식 (3.24)를 통해 식 (3.25)를 얻는다.

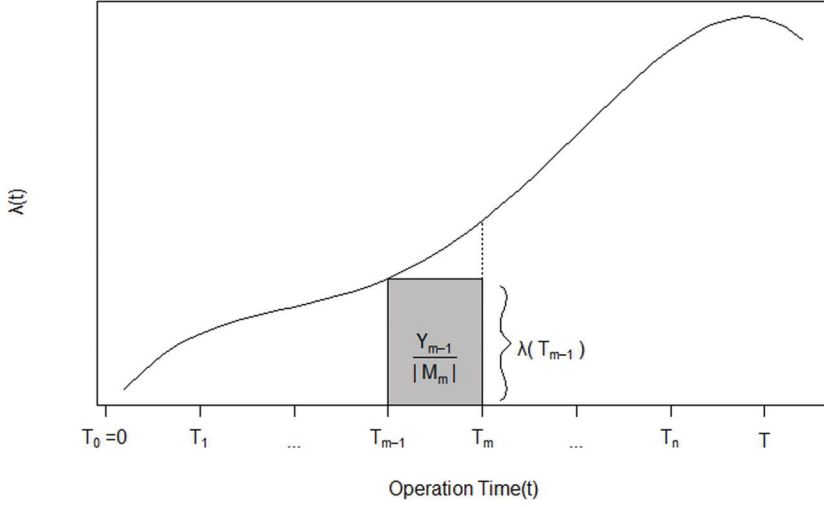


그림 12: 경험적 방법을 활용한 초기값 추정

$$\tilde{\lambda}(T_m) = \frac{Y_m}{|M_{m+1}|\Delta T} \quad (3.25)$$

결국 최종적으로 확인된 관계식인 식 (3.25)를 이용하여, 표 1.의 자료를 변환하면  $S = \{\tilde{\lambda}(T_m) : m = 0, 1, \dots, L-1\}$ 을 확보할 수 있으며, 이를 다시 식 (2.1)의 강도함수인  $\lambda(t) = \alpha\beta t^{\beta-1}$ 에 근사(fitting)하여 E-M 알고리즘의 초기값  $\alpha^{(0)}$ 와  $\beta^{(0)}$ 를 얻는다.

### 3.2.2 E-M 알고리즘의 활용

앞서 산출된  $\hat{\beta}$ 의 추정식인 식 (3.10)에서,  $r_j$ 는 알려져 있지 않기 때문에,  $r_j$ 는  $k$ 번째 E-step의 결과에서 산출된 값을 사용한다. 이 때,  $k$ 번째 E-step은  $(\alpha^{(k)}, \beta^{(k)})$ 와 관측값  $X_j = (s_j, x_{jr_j}, x_{j(r_j+1)}, \dots, x_{jn_j}, t_j)$ 을 이용하

여 획득된  $k$ 번째 M-step의 모수추정치를 활용한다. E-Step을 식 (3.1)의 형태로 기술하면 식 (3.26)과 같다. 특이점은,  $r_j$ 는 양의 정수이므로 기대값을 사용하는 것이 아닌 기대값을 만족하는 최대 정수를 사용해야 하며, 식 (3.26)에서  $\lfloor x \rfloor$ 는  $x$ 보다 크지 않은 양의 정수를 의미한다. 결국  $r_j$ 는  $s_j$  시점까지 발생된 누적 고장수이므로,

$$r_j^{(k+1)} = \left\lfloor E[N[0, s_j] | \alpha^{(k)}, \beta^{(k)}, X_j] \right\rfloor = \lfloor \alpha^{(k)} s_j^{\beta^{(k)}} \rfloor \quad k = 0, 1, 2, \dots \quad (3.26)$$

으로 표현된다. 또한, 관측된 고장수는  $(n_{j+1} - r_{j+1} + 1)$ 개이므로, 식 (3.10)의  $n_{j+1}$  역시 쉽게 구할 수 있다. 이후 산출된  $r_j^{(k+1)}$ 와  $n_j^{(k+1)}$ 를 이용하여  $\alpha^{(k)}$ 와  $\beta^{(k)}$ 를  $\alpha^{(k+1)}$  and  $\beta^{(k+1)}$ 로 최신화 하며, 이는 식 (3.9)와 식 (3.10)을 이용한다. 이상의 과정은 수렴조건

$$|\ln L(\alpha^{(k+1)}, \beta^{(k+1)}) - \ln L(\alpha^{(k)}, \beta^{(k)})| \leq \epsilon \quad (3.27)$$

을 만족할 때까지 반복되며, 최종적으로 산출된 모수는  $\hat{\alpha} = \alpha^{(k)}$ ,  $\hat{\beta} = \beta^{(k)}$ 가 된다.

### 3.2.3 신뢰구간 추정

추정된 모수들의 신뢰구간 추정에 앞서 추정값  $\hat{\alpha}, \hat{\beta}$ 의 분산(Variance)를 추정 하였다. 분산의 추정에는 Fisher Information 행렬,  $I$ 를 활용한다.

Schervish (2012); Nelson (2009)등에 의하면, ML 추정값,  $\theta$ 들은 일반적으로 식 (3.28)과 같은 관계를 갖기 때문에, Fisher Information 행렬의 역수는 구하고자 하는 모수의 공분산 행렬(Variance-covariance matrix)가



된다.

$$\hat{\theta} \sim N(\theta, I^{-1}(\theta)) \quad (3.28)$$

그러나 통상  $I(\theta)$ 는 알 수는 없으므로, 대표본을 전제로,  $I(\theta)$ 대신 Observed Fisher Information인 식(3.29)를 사용하는 것이 현실적이다. 실제로 이 같은 관계를 이용하여 Nelson (2009); Karbasian and Ibrahim (2010)에서 추정된 모수의 분산을 추정한 바 있다.

$$COV(\hat{\alpha}, \hat{\beta}) = \hat{I}^{-1} = - \begin{bmatrix} \frac{\partial^2 \ln \hat{L}}{\partial \alpha^2} & \frac{\partial^2 \ln \hat{L}}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ln \hat{L}}{\partial \alpha \partial \beta} & \frac{\partial^2 \ln \hat{L}}{\partial \beta^2} \end{bmatrix} = - \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (3.29)$$

식 (3.29)에서  $\frac{\partial^2 \ln \hat{L}}{\partial \alpha^2}$ ,  $\frac{\partial^2 \ln \hat{L}}{\partial \beta^2}$ ,  $\frac{\partial^2 \ln \hat{L}}{\partial \alpha \partial \beta}$ 는 Fisher information에서  $\alpha, \beta$ 를  $\hat{\alpha}, \hat{\beta}$ 로 대체한 것을 의미한다. 결과적으로 식 (3.28)에 따라, 식 (3.29)의 역행렬의 인자들인  $A, B, C$ 는 각각  $\text{Var}(\hat{\alpha})$ ,  $\text{Var}(\hat{\beta})$ ,  $\text{Cov}(\hat{\alpha}, \hat{\beta})$ 의 추정치와 같으며, 본 논문에서는 이 같은 방법으로 추정된 값들을  $\widehat{\text{Var}}(\hat{\alpha})$ ,  $\widehat{\text{Var}}(\hat{\beta})$ , and  $\widehat{\text{Cov}}(\hat{\alpha}, \hat{\beta})$ 로 표기한다. 이에 따라 유의수준  $\gamma$ 에서,  $\alpha, \beta$ 의 신뢰구간은 식 (3.30), (3.31)과 같다. 각 모수는 모형의 가정에 따라 음수가 될 수 없으므로, 이를 고려한 신뢰구간의 추정이 요구된다. 이 경우, 단순히 하한값이 음수가 아니도록 조정하는 방법과 모수의 로그 변환값의 신뢰구간을 역변환 하는 두 방법의 사용이 가능하다. 후자는 전자에 비해 하한값을 항상 0 이상인 값으로 산출해주는 장점이 있는 반면, 지수로 표현된 부분 때문에 분산의 크기가 클 수록 상·하한 값, 특히 상한값을 왜곡할 가능성이 존재한다. Power Law 과정의 모수는 상대적으로 작은 값들이며, 실제 3.3 절의 실험결과를 통해 두 방법을 비교 시 현재 방법이 더 나은 결과를 나타내고 있어, 본 연구에서는 하한값만 조정하는 방법을 사용하였다. ( $z_{\frac{\gamma}{2}}$  :

유의수준  $\frac{\gamma}{2}$ 에서의 정규분포 임계(critical)값)

$$\left( \max \left( 0, \hat{\alpha} - z_{\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\alpha})} \right), \hat{\alpha} + z_{\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\alpha})} \right) \quad (3.30)$$

$$\left( \max \left( 0, \hat{\beta} - z_{\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})} \right), \hat{\beta} + z_{\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})} \right) \quad (3.31)$$

식 (3.18), 식 (3.19), 식 (3.20)으로 제시된 신뢰성 관련 인자들은 아래와 같은 방법으로 분산과 신뢰구간을 추정하였다. 우선 고장률,  $\lambda(t)$ 는 항상 양수이므로, 로그를 취하여  $\alpha, \beta$ 에 관한 함수  $g(\alpha, \beta) = \ln \lambda(t)$ 인  $g(\alpha, \beta)$ 를 정의한다면, 식 (3.32)와 같은 임의의 지점  $(\alpha_0, \beta_0)$ 에 대한 테일러 전개 (Taylor Expansion)를 통하여  $\lambda(t)$ 의 분산을 얻을 수 있다.

$$\begin{aligned} g(\alpha, \beta) &\approx g(\alpha_0, \beta_0) + (\alpha - \alpha_0) \times \frac{\partial g}{\partial \alpha} \Big|_{\alpha=\alpha_0} + (\beta - \beta_0) \times \frac{\partial g}{\partial \beta} \Big|_{\beta=\beta_0} \\ &= \text{const} + (\alpha - \alpha_0) \left( \frac{1}{\alpha_0} \right) + (\beta - \beta_0) \left( \frac{1}{\beta_0} + \ln t \right) \end{aligned} \quad (3.32)$$

$g(\alpha, \beta)$ 의 분산,  $\widehat{\text{Var}}(g(\hat{\alpha}, \hat{\beta}))$ 은 분산의 성질에 의해 식 (3.33)으로 표현되고,

$$\begin{aligned} \widehat{\text{Var}}(g(\hat{\alpha}, \hat{\beta})) &\approx \text{Var} \left( (\hat{\alpha} - \alpha_0) \left( \frac{1}{\alpha_0} \right) + (\hat{\beta} - \beta_0) \left( \frac{1}{\beta_0} + \ln t \right) \right) \\ &= \left( \frac{1}{\alpha_0} \right)^2 \text{Var}(\hat{\alpha}) + \left( \frac{1}{\beta_0} + \ln t \right)^2 \text{Var}(\hat{\beta}) \\ &\quad + \left( \frac{2}{\hat{\alpha}} \right) \left( \frac{1}{\hat{\beta}} + \ln t \right) \text{Cov}(\hat{\alpha}, \hat{\beta}) \end{aligned} \quad (3.33)$$

최종적으로  $(\alpha_0, \beta_0)$ 을  $(\hat{\alpha}, \hat{\beta})$ 로 대체하면  $\widehat{\text{Var}}(g(\hat{\alpha}, \hat{\beta}))$ 은 식 (3.34)의 형태

가 된다.

$$\begin{aligned}\widehat{\text{Var}}(g(\hat{\alpha}, \hat{\beta})) &\approx \left(\frac{1}{\hat{\alpha}}\right)^2 \widehat{\text{Var}}(\hat{\alpha}) + \left(\frac{1}{\hat{\beta}} + \ln t\right)^2 \widehat{\text{Var}}(\hat{\beta}) + \left(\frac{2}{\hat{\alpha}}\right) \left(\frac{1}{\hat{\beta}} + \ln t\right) \widehat{\text{Cov}}(\hat{\alpha}, \hat{\beta}) \\ &\equiv V_1\end{aligned}\quad (3.34)$$

위와 같은 방법으로 산출된  $V_1$ 은  $\ln \lambda(t)$ 의 분산이므로, 실제 유의수준  $\gamma$ 에서  $\lambda(t)$ 의 신뢰구간은

$$\left( \hat{\lambda}(t) \exp\left(-z_{\frac{\gamma}{2}} \sqrt{V_1}\right), \hat{\lambda}(t) \exp\left(z_{\frac{\gamma}{2}} \sqrt{V_1}\right) \right) \quad (3.35)$$

로, 식 (3.20)에서  $MTBF(t)$ 의 신뢰구간은

$$\left( \frac{1}{\hat{\lambda}(t)} \exp\left(-z_{\frac{\gamma}{2}} \sqrt{V_1}\right), \frac{1}{\hat{\lambda}(t)} \exp\left(z_{\frac{\gamma}{2}} \sqrt{V_1}\right) \right) \quad (3.36)$$

으로 표현된다. 또한, 식 (3.32)와 유사한 방법을 사용한  $m(t)$ 의 신뢰구간은

$$\left( \hat{m}(t) \exp\left(-z_{\frac{\gamma}{2}} \sqrt{V_2}\right), \hat{m}(t) \exp\left(z_{\frac{\gamma}{2}} \sqrt{V_2}\right) \right) \quad (3.37)$$

과 같이 구할 수 있다. 이 때,  $V_2$ 는 식 (3.38)을 의미한다.

$$V_2 \equiv \left(\frac{1}{\hat{\alpha}}\right)^2 \widehat{\text{Var}}(\hat{\alpha}) + (\ln t)^2 \widehat{\text{Var}}(\hat{\beta}) + \left(\frac{2 \ln t}{\hat{\alpha}}\right) \widehat{\text{Cov}}(\hat{\alpha}, \hat{\beta}) \quad (3.38)$$

### 3.3 제안된 모형의 성능 확인

본 장에서는 앞서 제시된 수학적 모형을 다양한 모의실험(Simulation) 조건들에 대해 적용하여 제안된 모형의 타당성 확인과 적용 범위 등을 확인한다. 실험 방법은 다음과 같다. 우선 동일한 Power Law 과정을 따

르는 50개의 독립된 장비에서 50에서 100개 사이의 고장을 임의 발생시킨다. 이후 자료 손실상황을 모사하기 위해 정해진 손실비율(Missing Ratio,  $MR$ )에 따라 장비별 고장시점 중  $[0, MR]$ 구간의 자료를 배제 한 후 나머지 자료를 실험에 사용한다. 각 실험 조건에 대한 random화의 영향을 줄이기 위해 100번씩 반복 수행하였고, 모의실험 코드 구현에는 R Code(TEAM (2010))를 활용하였다. 모의실험 조건은 표 2.와 같다.

표 2: 기본모형의 성능확인을 위한 실험조건

Cases		Level
장비수 (대)	50	6
장비당 고장수 (개)	50 to 100 (randomly)	
손실수준, $MR(\%)$	1, 2, 5, 10, 20, 30	
$\alpha$	0.05 to 0.5 by 0.02	
$\beta$	0.5 to 3.0 by 0.05	
반복수 (회)	100	26
Total		358,800

### 3.3.1 모형의 효용성과 민감도

앞서 제안한 방법은 좌측절단이 발생한 자료만을 사용하여도 완전한 자료를 사용했을 때와 유사한 추정결과를 나타낼 수 있을 때 의미를 갖는다. 또한 E-M 알고리즘 사용의 타당성은 경험적 방법에 의해 산출된 초기값  $\alpha^{(0)}, \beta^{(0)}$ 을 충분히 개선할 수 있음에서 확인할 수 있다. 모의실험의 효과 측정 지표로는 NRMSE(Normalized Root Mean Squared Error)를 사용하였다. 본 논문에서 사용한 NRMSE는 관측값 평균을 표준화 기준으로 사용하였으며, 관측값이  $N$ 개이고 참값이  $\theta$ , 추정치가  $\theta_i, i = 1 \dots N$ 일

때 식 (3.39)로 표현된다.

$$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{\theta - \hat{\theta}_i}{\theta} \right)^2} \quad (3.39)$$

NRMSE는 RMSE를 표준화한 값으로, 비교대상의 크기에 관계없이 비교가 가능한 지표이기 때문에 크기의 차이가 있는 모수의 효율을 비교하기 위해 채택하였다. NRMSE는 CV(Coefficient of variation of RMSE)라는 명칭으로도 사용되며, 알려진 참값이나 관측값(Bouman and Van Laar (2006)), 관측값 범위(Nicolardot et al. (2001))를 표준화 기준으로 한다. 경우에 따라 표본 RMSE(Leiva et al. (2016))나 표준편차(Chan and Cannon (2002))를 표준화 기준으로 사용하기도 한다. 하지만 본 논문에서는 이미 참값을 알고 있고 단순히 참값과의 오차를 비율로 환산하는 것이 목적이므로 표준화는 모의실험에 사용된 참값을 기준으로 사용하였다.

우선 좌측절단이 발생한 자료로 모수를 추정한 결과를 그림 13., 그림 14.와 같다. 좌측이  $\alpha$ , 우측이  $\beta$ 의 NRMSE를 의미하며 NRMSE는 모수추정 대상이 되는 참값  $\alpha, \beta$  조합에 따라 2차원 등고선 형태 그래프로 나타내었다. 오차의 크기는 음영으로 표시되며, 음영이 짙어질수록 참값과의 오차는 커짐을 의미한다.

손실자료가 많을수록 모수추정값의 오류는 당연히 증가하는 것이 당연한 현상이다. 다만, 그 정도는  $\alpha$ 가  $\beta$ 에 비해 더 뚜렷함을 알 수 있다. 우선 10% 이하의 손실비율에서는 모든 조합에 대해 오차의 분포에 관한 뚜렷한 패턴이 드러나지 않으나, 그 이상의 경우 오차의 분포에 경향성이 발견되기 시작한다.  $\alpha$ 의 경우 실제  $\beta$ 가 1.0 이하이고 실제  $\alpha$ 가 0.15 부근일 경우 NRMSE가 커지는 경향을 나타내며,  $\beta$ 의 경우 모든 조합에서 실제  $\alpha$ 가 0.15 부근일 경우 NRMSE가 큰 것을 알 수 있다. 즉,  $\alpha$ 가 0.15 부근으로



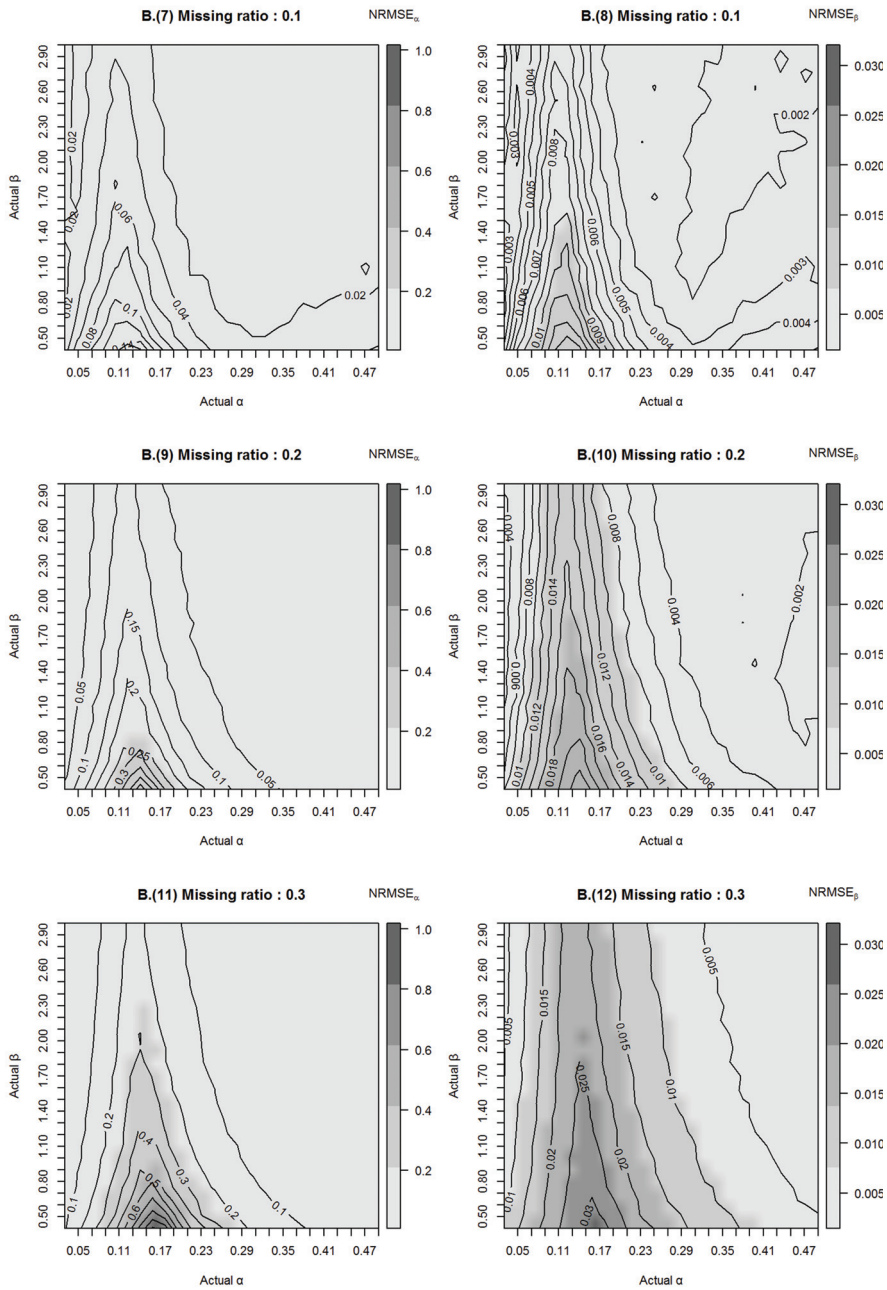


그림 14: 모의실험 결과 NRMSE(MR : 0.1, 0.2, 0.3)

추정되는 경우 다른 추정값에 비해 유의하여 사용해야 한다는 결론을 얻을 수 있다. 그림 13.와 그림 14.에서 나타나는 경향성을 좀 더 뚜렷하게 확인하기 위해 손실 비율에 따라 발생된 오차를 표 3.으로 재정리 하였다. 표 3.에 제시된 평균 오차는 각각의 손실비율에 대한 모의조건 598개 ( $= 23 \times 26$ )에서 산출된 NRMSE 평균이다. 이를 통해 다음과 같은 사실을 확인하였다.

표 3: 기본모형의 손실비율에 따른 추정효율( $\overline{NRMSE}$ )

손실수준 (%)	$\alpha^{(0)}$	$\hat{\alpha}$	$\beta^{(0)}$	$\hat{\beta}$
1.0	2.0220	0.0164	0.0346	0.0027
2.0	1.9607	0.0165	0.0347	0.0028
5.0	1.8168	0.0185	0.0348	0.0030
10.0	1.7102	0.0261	0.0352	0.0039
20.0	1.8237	0.0590	0.0370	0.0069
30.0	2.0582	0.1255	0.0386	0.0114

오차의 크기는 항상  $\alpha$ 가  $\beta$ 에 비해 10배가량 크게 나타난다. 실험에 사용된 조건 중 최악조건인 30% 손실에서  $\alpha$ 는 무려 13% 가량의 오차를 나타내나,  $\beta$ 는 1%남짓 한 오차만을 나타내고 있다. 이상의 결과로부터 모형의 정확도와 제한점을 확인하여, 활용 영역을 다음의 두 영역으로 요약할 수 있었다.

우선, 표 3.으로부터  $\beta$ 의 오차는 최대 2%를 넘지 않음을 알 수 있다. 이는 매우 유의미한 결과이다. 왜냐하면, Power Law 과정을 통해 확인하고자 하는 가장 핵심적인 결과인 “고장의 시간에 따른 변동”을 표시하는 것이  $\beta$ 이기 때문이다. 제안된 모형에 의해서 다량의 자료가 유실된 경우에도  $\beta$ 는 1% 가량의 오차만을 나타내며 추정이 가능하기 때문에, 고장의 경향성은 충분히 분석할 수 있다. 즉, 제안된 모형은 자료의 손실이 많더



라도 정성적인 추세분석에는 사용이 가능하다.

다음으로, 특정시점에서의 MTBF를 산출하거나 미래의 고장시점을 예측하는 등 정량적 분석이 가능한 최대 손실비율 수준을 확인할 수 있었다. 비록  $\alpha$ 의 오차가  $\beta$  오차에 비해 항상 크지만 특정 시점까지는 오차의 증가량이 크지 않음을 확인할 수 있다. 그림 15.는 손실 비율별 오차의 크기를 나타낸 것으로  $\beta$ 의 오차는 큰 변화가 없지만 10% 이상의 손실이 발생될 경우  $\alpha$ 의 오차는 크게 변동된다.  $\alpha$ 의 오차가 1% 미만의 손실 상황에서 1.6%가량 발생하는 것을 고려하면 10% 수준의 손실이 발생될 때까지는 큰 무리 없이 활용 가능하다는 결론을 얻는다.

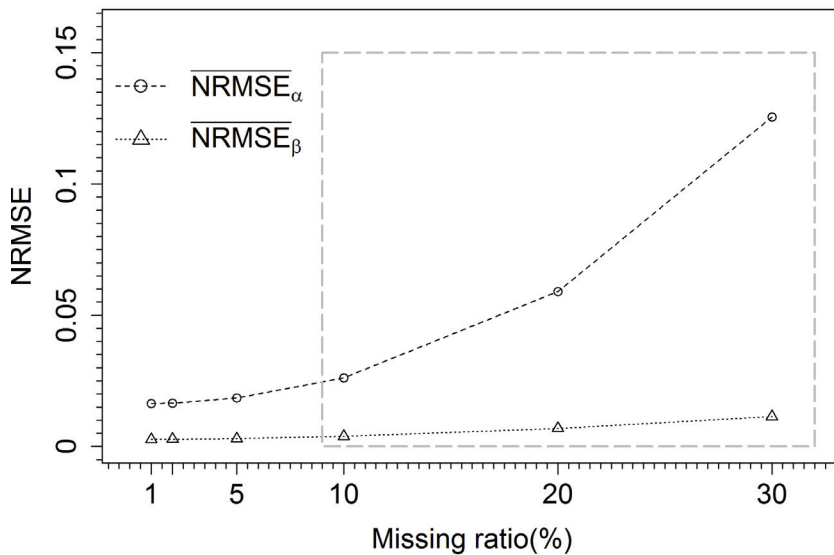


그림 15: 기본모형의 손실비율에 따른 추정효율( $\overline{NRMSE}$ )

다시 말하면, 10% 이하의 손실 상황까지는 모수추정에 큰 영향을 주지 않기 때문에, 이 범위 내에서 정량적 분석이 가능하다고 결론지을 수 있

다. 표 4.는 정량적 분석이 가능한 손실수준의 평균인 5% 손실 상황에서 모의된 모수 조합의 극단값 4개의 추정치와 신뢰구간을 나타낸 것이다. 표 4.로부터, 추정값과 실제값은 유사하며 99% 신뢰구간 내에 실제값을 포함하는 등, 제시된 정량적 분석범위가 타당함을 간접적으로 확인할 수 있다.

표 4: 기본모형의 모수추정 사례 (5% 손실수준)

$\alpha$	$\beta$	$\hat{\alpha}$	표준편차	$\hat{\beta}$	표준편차
0.05	0.5	0.0576	0.0092	0.4856	0.0112
0.05	3.0	0.0702	0.0090	2.8523	0.0567
0.49	0.5	0.3513	0.0441	0.5269	0.0123
0.49	3.0	0.5025	0.0403	2.9780	0.0523

다음으로 초기값의 개선효과 관련 사항이다. 표 3.에서 확인할 수 있듯이 초기값  $\alpha^{(0)}, \beta^{(0)}$ 의 평균오차는 E-M 알고리즘에 의해 개선된 추정값  $\hat{\alpha}, \hat{\beta}$ 의 평균오차보다 항상 큰 값을 나타내고 있다. 이를 통해 E-M 알고리즘의 효과성은 충분히 확인할 수 있었다.

### 3.3.2 신뢰구간의 정확성

본 절에서는 앞서 제시된 모수추정치와 신뢰도 관련 지표에 대한 신뢰구간의 타당성을 확인한다. 앞서 민감도 분석에서 사용했던 방법과 동일하게  $\alpha = 0.1$  and  $\beta = 1.2$ 이고, 손실수준 5%에서 모의실험을 수행하였고, 결과는 표 5.와 같다.

표 5.로부터 E-M 알고리즘의 적용은 초기해  $\alpha^{(0)}, \beta^{(0)}$ 를 충분히 개선하고 있음을 알 수 있으며, 최종 추정값의 오차인  $|\frac{\hat{\alpha}-\alpha}{\alpha}|$ 와  $|\frac{\hat{\beta}-\beta}{\beta}|$ 가 각각 5.4%, 0.97%에 불과함을 알 수 있다. 이는 앞서 제시된 결과와 일치하는 결과이기도 하다. 추가적으로 모의에 사용된 참값이 추정값의 95%, 99%

표 5: 기본모형의 신뢰구간 산식의 타당성 확인결과

	$\alpha$	$\beta$
실제모수	0.1	1.2
초기값 ( $\alpha_0, \beta_0$ )	0.2668	1.0851
추정 결과 ( $\hat{\alpha}, \hat{\beta}$ )	0.1054	1.1884
오차 ( $\hat{\alpha}, \hat{\beta}$ ) (%)	5.4	0.97
95% 신뢰구간	[0.0842, 0.1266]	[1.1534, 1.2234]
99% 신뢰구간	[0.0776, 0.1333]	[1.1424, 1.2344]

신뢰구간 내에 충분히 위치하고 있음도 확인된다.

$t$ 의 변화에 따른  $\widehat{MTBF}(t)$ 의 신뢰구간은 그림 16.으로 표현된다.

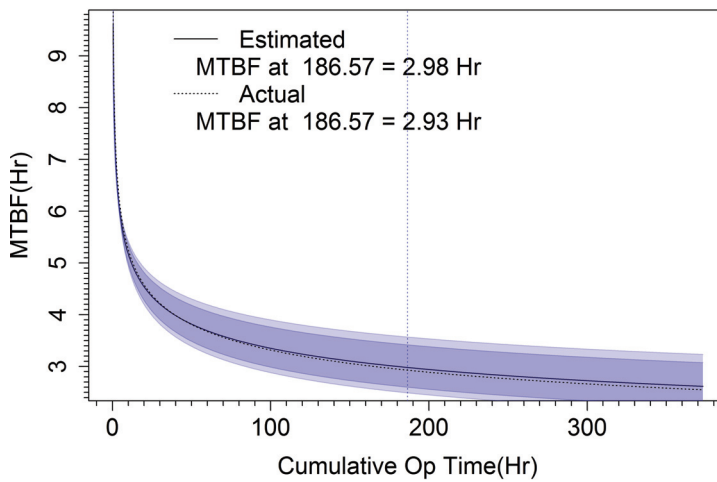


그림 16:  $\widehat{MTBF}(t)$ 와 신뢰구간

그림 16.에서, 시간에 따라 변화되는 MTBF들이  $\widehat{MTBF}(t)$ 의 신뢰구간 내에 위치함이 확인된다. 모의에 사용된 시간 평균인 186.57시간에서  $MTBF(186.57) = 2.93$ 시간으로 확인되며, 추정치  $\widehat{MTBF}(t) = 2.98$ 시간으

로 근소한 차이만 발생됨을 알 수 있다.

$t$ 의 변화에 따른  $m(t)$ 의 변화 역시 유사한 결과를 보인다. 그림 17.은  $m(t)$ 를 신뢰구간과 함께 나타낸 것이다.

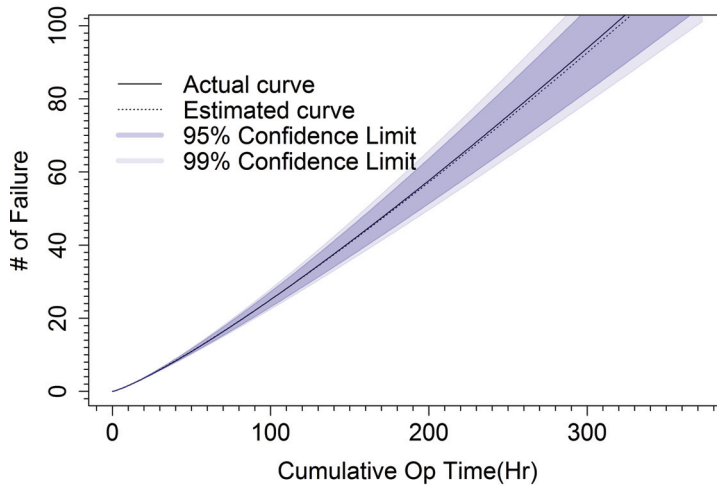


그림 17:  $\widehat{m(t)}$ 와 신뢰구간

흑색 실선으로 표시된 참값과 흑색 점선으로 표시된 추정값은 거의 일치하며, 모든 시점에서 신뢰구간 내에 위치하는 것으로 분석되고 있다. 이상의 결과로부터, 3.2.3절에 제시된  $\widehat{MTBF(t)}$ 와  $\widehat{m(t)}$ 의 신뢰구간 산정식의 유효성을 확인할 수 있었으며,  $\widehat{MTBF(t)}$ 와  $\widehat{m(t)}$ 의 추정방법의 적절성 역시 확인할 수 있었다.

## 제 4 장

### 이질요소를 고려한 모형의 확장

군용장비와 같이 정해진 작전운용개념이 존재하고 대부분 유사한 환경에서 운영되는 경우, 장비들 간의 이질성을 고려하지 않더라도 큰 문제는 없을 수 있다. 그러나 실제로 동일한 장비라 하더라도 운용자 개개인의 습관이나 노출되는 환경에 따라 고장의 발생빈도는 다소 상이할 수 있다. 이러한 동질성을 저해하는 요소에는 사용빈도, 노출온도, 사용 군 등 기록이나 식별이 가능한 요소가 있는 반면, 연도별 제작기술 변화, 사용자 습관 등 상이한 특성을 나타낼 것으로 예상되나 기록으로 구분하기 어려운 요소도 있다. 흔히 전자는 공변량(Covariate)으로 식별하여 모형에 반영하는 것이 일반적이며, 후자는 Frailty라는 요소로 고려되곤 한다. 본 논문에서는 위 두 가지 요소를 기본모형의 강도함수에 점진적으로 반영하는 방식으로 모형 확장을 시도하였다. 우선 기본모형에 Frailty를 반영하여 1차로 모형을 확장하고, 여기에 2차로 공변량을 추가하는 방법을 사용한다. 이는 2.3.3절에서 언급한 바와 같이 선행 연구결과를 통해 입증된 모형의 특성을 활용키 위한 것이기도 하며, 결국 두 요소 모두 추가된 모형으로 어느 하나의 요소만 보유한 자료 역시 분석 가능하기 때문이다. 전자와 관련하여 4.2.1절에서 다시 언급되지만, 공변량 추가 모형의 경우, 약간의 변형을 통해 Frailty 추가모형의 모수와 Frailty 규모 추정식을 재활용하기 때문에 Frailty만을 추가한 모형을 우선 제시한다.

## 4.1 관측 불가능한 이질요소(Frailty) 반영

### 4.1.1 강도함수의 모수추정

Frailty를 확률모형에서 고려하고자 하는 노력은 새로운 것은 아니다. Vaupel et al. (1979) 이후로 Bhattacharjee et al. (2002); Cha and Finkelstein (2014); Slimacek and Lindqvist (2013)등이 다양한 방법으로 고장추이 분석과 관련하여 Frailty를 고려하고자 시도하였으며, 모형의 추가변수로 고려하거나 기존 모형의 모수에 베이지안 기법을 적용하는 방법을 사용해 왔다. 본 논문의 기본모형에서 사용된 강도함수에 Frailty를 반영한 Asfaw and Lindqvist (2015)의 연구 역시 Frailty,  $Z$ 가 감마분포( $\Gamma(k, \theta)$ ) 임을 가정하고 있다. 통상, Frailty는 Frailty의 크기보다는 분산을 추정하는 것이 목적이고,  $E(Z) = 1$ ,  $\text{Var}(Z) = \delta$ 인 가정을 하기 때문에,  $Z$ 의 확률밀도함수는 식 (4.1)의 형태가 된다.

$$h(z) = \frac{z^{1/\delta-1} \text{Exp}[-z/\delta]}{\delta^{1/\delta} \Gamma(\frac{1}{\delta})} \quad (4.1)$$

문제는 Frailty자체가 관측이 불가능한 것임에도 불구하고 특정 확률분포를 따른다는 가정이 무리가 있다는 사실이다. Slimacek and Lindqvist (2016)는 마치 베이지안 모형에서 사전확률분포를 선택하는 과정에 비유하기도 하였다. 실제로 Asfaw and Lindqvist (2015)의 연구에서도 Frailty를 감마분포로 가정하였지만, 실제로 감마분포인지 아닌지는 알 수 없고 다만, 감마분포 사용이 실제 자료와 잘 맞고 수학적 표현이 용이하기 때문에 사용하였다고 인정하였다. 또 다른 문제점으로는 모형이 불필요하게 복잡해진다는 점에 있다. 결국 모수추정의 대상은 식 (4.1)에서  $z$ 가 아니라  $\alpha, \beta, \delta$ 가 되기 때문에, 실제  $z$ 가 포함된 우도함수,  $L(z)$ 를 식 (4.2)와 같이

integrating out한 우도함수를 사용하게 된다. 이 과정은 대부분 산식이 복잡하기 때문에 수치해석을 필요로 한다.

$$L = E_z[L(z)] = \int L(z)h(z)dz \quad (4.2)$$

위와 같은 단점을 극복하기 위해 Slimacek and Lindqvist (2016)는 Frailty를 비모수적으로 처리하는 방법을 제안하였다. 개별 장비의  $z_j$ 들은 분포를 가지는 확률변수가 아닌 추정이 필요한 모수들로 처리하는 방식이다. 각  $z_j$ 들은 상호 독립이며, 확률변수  $Z$ 의 실현값(realization)으로 Asfaw and Lindqvist (2015)의 가정과 마찬가지로  $E(Z) = 1$ 을 가정한다. 이는 일반적인 회귀분석이나 고정효과(fixed effect) 모형과 유사하다. Slimacek and Lindqvist (2016)의 방법은 불필요한 분포가정을 하지 않는다는 점에서 좀 더 계산이 간편하고 직관적인 장점이 있으며, 관측중단시점( $t_j$ )의 차이가 크지 않다면 분포가정을 하는 기존 논문만큼의 효율을 나타낼 수 있는 것으로 분석되었다. Slimacek and Lindqvist (2016)의 모형을 활용하여 기존 논문의 강도함수를 식 (4.3)으로 변형하고 좌측절단 상황에 부합하도록 수정을 시도하였다.

$$\lambda(t) = \alpha_j \beta t^{\beta-1} \quad (4.3)$$

기존논문의 강도함수 중 척도모수  $\alpha$ 를  $z_j\alpha$ 가 아닌  $\alpha_j$ 로 변형한 것은, 전자로 변환하였을 때 식별문제(identification problem)이 발생되기 때문이다 (Slimacek and Lindqvist (2016)). 이 때문에  $z_j\alpha$ 가 아닌 동일한 성격의 다른 변수  $\alpha_j$ 를 강도함수의 모수로 사용한다.

식 (4.3)을 Power Law 과정의 강도함수로 사용할 때 부록 1.과 유사

한 방법으로 우도함수를 산출하여 식 (4.4)를 얻을 수 있다. 식 (4.4)에서  $H = \sum_{j=1}^N \sum_{k=r_j}^{n_j} \ln x_{j,k}$ ,  $K = \sum_{j=1}^N (n_j - r_j + 1)$ 를 의미한다.

$$\begin{aligned} \ln L \propto & \sum_{j=1}^N n_j \ln \alpha_j + K \ln \beta \\ & + \beta \sum_{j=1}^N (r_j - 1) \ln s_j + (\beta - 1)H - \sum_{j=1}^N \alpha_j t_j^\beta \end{aligned} \quad (4.4)$$

식 (4.4)의 score equation인 식 (4.5)에서

$$\begin{aligned} \frac{\partial \ln L}{\partial \alpha_j} &= \frac{n_j}{\alpha_j} - t_j^\beta = 0 \\ \frac{\partial \ln L}{\partial \beta} &= \frac{K}{\beta} + \sum_{j=1}^N (r_j - 1) \ln s_j + H - \sum_{j=1}^N \alpha_j t_j^\beta \ln t_j = 0 \end{aligned} \quad (4.5)$$

$\hat{\alpha}_j, \hat{\beta}$ 는 결국 식 (4.6)으로 추정된다.

$$\begin{aligned} \hat{\alpha}_j &= \frac{n_j}{t_j^\beta} \\ \hat{\beta} &= \frac{K}{\sum_{j=1}^N n_j \ln t_j - \sum_{j=1}^N (r_j - 1) \ln s_j - H} \end{aligned} \quad (4.6)$$

식 (4.6)에 제시된 추정결과는 III장에 제시된 기본모형과 달리  $\beta$ 의 추정식이 closed-form으로 산출되어 추정에 별도의 수치 해석적 방법의 사용은 불필요한 것으로 확인된다. 추가로 반영된 인자인  $\hat{z}_j$ 를 추정은 Slimacek and Lindqvist (2016)에 제시된  $\alpha$ 와  $n t^\beta$ 간의 관계를 활용하여 산출한다.



$\hat{z}_j$ 에 사용된 NHPP의 성질은 다음과 같다. 우선 강도함수가 식 (2.7)인 NHPP에서 총 고장 발생 횟수,  $n$ 은 임의의  $Z, t$  조건에서, NHPP의 기본 성질로부터

$$E(n|Z, t) = Z\alpha t^\beta \quad (4.7)$$

이 됨을 알 수 있다. 위 관계는 조건부 기대값이므로, 양 변에 기대값을 취하여 조건을 제거하면,

$$E(n) = E(E(n|Z, t)) = \alpha E(t^\beta) \quad (4.8)$$

가 된다. 위 식에서 양변을  $E(t^\beta)$ 로 나누면,  $\alpha$ 는

$$\alpha = \frac{E(n)}{E(t^\beta)} \quad (4.9)$$

임을 알 수 있다. 산식의 활용을 위해  $\frac{n}{t^\beta}$ 에 대해서도 유사한 방법을 적용하면,

$$\begin{aligned} E\left(\frac{n}{t^\beta}\right) &= E\left(E\left(\frac{n}{t^\beta} | Z, t\right)\right) \\ &= E\left(\frac{1}{t^\beta} E(n|Z, t)\right) \\ &= \alpha \end{aligned} \quad (4.10)$$

이 된다. 결국, 식 (4.8)과 식 (4.10)으로부터, 식 (4.11)과 같은  $\alpha$ 와  $n, t^\beta$ 의 관계식을 얻을 수 있다. 이는 향후 식 (4.6)으로 추정된  $\alpha_j$ 들에서  $z_j$ 와  $\alpha$ 를

분해하는 기본 성질로 활용된다.

$$\alpha = \frac{E(n)}{E(t^\beta)} = E\left(\frac{n}{t^\beta}\right) \quad (4.11)$$

그런데,  $\frac{n}{t^\beta}$ 은  $\alpha_j$ 의 ML 추정치이므로, 결국 식 (4.11)의 sample form으로부터,  $\hat{\alpha}$ 은

$$\hat{\alpha} \equiv \frac{1}{N} \sum_{j=1}^N \hat{\alpha}_j \quad (4.12)$$

이 되고, 앞서 가정한  $\alpha_j = z_j \alpha$ 에 의해,  $\hat{z}_j$ 는

$$\begin{aligned} \hat{z}_j &= \frac{\hat{\alpha}_j}{\hat{\alpha}} = \frac{n_j/t_j^{\hat{\beta}}}{\hat{\alpha}} \\ &= \frac{n_j/t_j^{\hat{\beta}}}{(1/N) \sum_{j=1}^N (n_j/t_j^{\hat{\beta}})}, \quad j = 1, \dots, N \end{aligned} \quad (4.13)$$

과 같이 추정된다. 식 (4.6)에서, 실제로는  $r_j$ 를 알 수 없기 때문에  $r_j$ 와  $\alpha_j$ ,  $\beta$ 의 추정에 E-M 알고리즘을 사용하게 된다. 수정된 모수추정 절차는 그림 18.과 같다. 다만, Frailty 고려모형은 기본모형의 확장이므로 초기값 산출은 3.2.1절의 방법을 우선 사용한다.

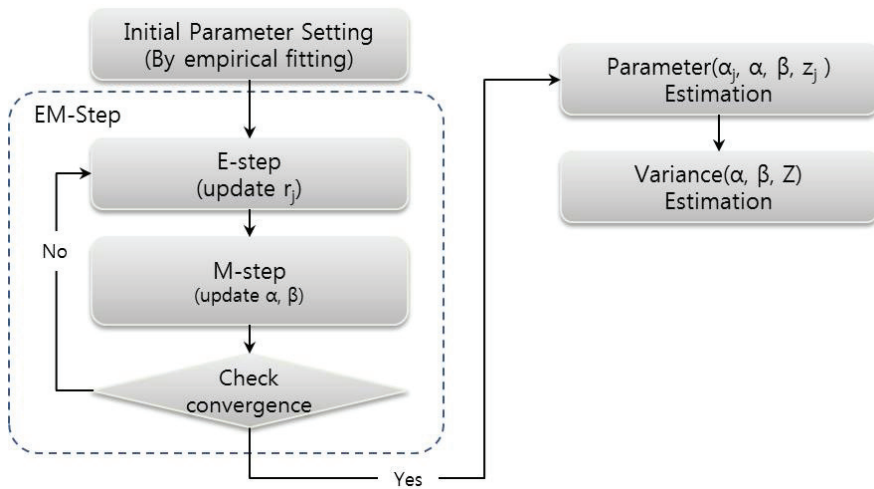


그림 18: Frailty를 고려한 확장모형의 모수추정 절차

#### 4.1.2 모수의 분산과 Frailty의 크기 추정

확장모형 역시 기본모형에서 사용했던 것과 동일한 방법으로 관측된 Fisher Information 행렬을 활용하여 모수의 분산  $\widehat{\text{Var}}(\hat{\alpha})$ ,  $\widehat{\text{Var}}(\hat{\beta})$ 와 Frailty의 크기, 즉  $Z$ 의 분산  $\widehat{\text{Var}}(\hat{Z})$ 를 추정한다. Fisher Information,  $\hat{I}$ 는  $N$ 개의  $\alpha_j$ 와 1개의  $\beta$ 가 있으므로  $(N+1) \times (N+1)$  행렬로 표현된다.  $I$ 의  $i$ 행  $j$ 열

인자를  $F_{ij}$ 라 할 때, 각 인자는 식 (4.14)에서 식 (4.18)과 같다.

$$i = j \neq (N+1) :$$

$$F_{ij} = \frac{n_j}{\hat{\alpha}_j^2} \quad (4.14)$$

$$i = j = (N+1) :$$

$$F_{(N+1),(N+1)} = \frac{K}{\hat{\beta}^2} + \sum_{j=1}^N \hat{\alpha}_j t_j^{\hat{\beta}} (\ln t_j)^2 \quad (4.15)$$

$$i \neq j < (N+1) :$$

$$F_{ij} = 0 \quad (4.16)$$

$$i = (N+1), i < (N+1) :$$

$$F_{(N+1),j} = t_j^{\beta} \ln t_j \quad (4.17)$$

$$j = (N+1), i < (N+1) :$$

$$F_{i,(N+1)} = t_i^{\beta} \ln t_i \quad (4.18)$$

위 식들을 각 인자로 갖는 Fisher Information의 역행렬은 Schott (2016)의 블록 행렬(Partitioned Matrix)의 역행렬과 관련된 성질을 이용하여 계산할 수 있고, 이를 활용한  $\widehat{\text{Var}}(\hat{\beta}|Z, t)$ ,  $\widehat{\text{Var}}(\hat{\alpha}_j|Z, t)$ 와  $\widehat{\text{Var}}(\hat{\alpha}_i, \hat{\alpha}_j|Z, t)$ 의 추정결과는 식 (4.19), 식 (4.20), 식(4.21)과 같다. 세부 과정은 Appendix A.에 별도로 수록하였다.

$$\begin{aligned} \widehat{\text{Var}}(\hat{\alpha}_j|Z, t) &= \frac{\hat{\alpha}_j^2}{n_j} + \frac{\hat{\beta}^2}{K} \hat{\alpha}_j^2 (\ln t_j)^2 \\ &= \frac{\hat{\alpha}_j^2}{n_j} + \widehat{\text{Var}}(\hat{\beta}) \hat{\alpha}_j^2 (\ln t_j)^2 \end{aligned} \quad (4.19)$$

$$\begin{aligned}
\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\alpha}_j | Z, t) &= \frac{\hat{\beta}^2}{K} \hat{\alpha}_i(\ln t_i) \hat{\alpha}_j(\ln t_j) \\
&= \widehat{\text{Var}(\hat{\beta})} \hat{\alpha}_i \hat{\alpha}_j (\ln t_i) (\ln t_j)
\end{aligned} \tag{4.20}$$

$$\widehat{\text{Var}}(\hat{\beta} | Z, t) = \frac{\hat{\beta}^2}{K} \tag{4.21}$$

위 결과로부터, 실제  $\hat{\alpha}$ 는  $\hat{\beta}$ 로 표현할 수 있으므로, 전체적인 Fisher information은  $\hat{\beta}$ 로만 표현됨을 알 수 있다. 또한, 위 식에서,  $\text{Var}(\hat{\beta})$ 는  $Z, t$ 와 무관한 unconditional 값이지만,  $\text{Var}(\hat{\alpha})$ 는  $Z$ 를 전제로 한 조건부 값이고,  $\alpha$ 가 아닌  $\alpha_j$ 들에 관한 식임을 알 수 있다. 따라서  $\text{Var}(\hat{\alpha}_j)$ 의 조건을 제거하고 식 (4.12)의  $\alpha_j$ 와  $\alpha$  관계식, 분산의 성질  $\text{Var}(A) = E[\text{Var}(A|B)] + \text{Var}[E(A|B)]$ 을 통해  $\text{Var}(\hat{\alpha})$ 를 산출한다(Slimacek and Lindqvist (2016)). 결과는 식 (4.22)와 같다.

$$\begin{aligned}
\text{Var}(\hat{\alpha}) &= E(\text{Var}(\hat{\alpha} | Z, t)) + \text{Var}(E(\hat{\alpha} | Z, t)) \\
&\approx \left[ \frac{1}{N^2} \sum_{j=1}^N \text{Var}(\hat{\alpha}_j | Z, t) + \frac{2}{N^2} \sum_{i \leq j} \text{Cov}(\hat{\alpha}_i, \hat{\alpha}_j | Z, t) \right] \\
&\quad + \left[ \hat{\alpha}^2 \frac{\text{Var}(Z)}{N} \right]
\end{aligned} \tag{4.22}$$

그런데, 식 (4.22)는  $\text{Var}(Z)$ 를 포함하고 있으므로,  $\text{Var}(\hat{\alpha})$ 의 산출을 위해서는 우선  $\text{Var}(Z)$ 를 산출해야 한다.  $\text{Var}(Z)$  역시 NHPP의 기본 성질과 조건부 분산식을 이용하여 산출한다. 아래 식 (4.23), 식 (4.24)를 식 (4.25)

에 반영하면,

$$E\left(\text{Var}(n|Z, t)\right) = \alpha E(t^\beta) \quad (4.23)$$

$$\text{Var}\left(E(n|Z, t)\right) = E(Z^2)\alpha^2 E(t^{2\beta}) - \alpha^2 E(t^\beta)^2 \quad (4.24)$$

$$\begin{aligned} \text{Var}(n) &= E\left(\text{Var}(n|Z, t)\right) + \text{Var}\left(E(n|Z, t)\right) \\ &= \alpha E(t^\beta) + E(Z^2)\alpha^2 E(t^{2\beta}) - \alpha^2 E(t^\beta)^2 \end{aligned} \quad (4.25)$$

이 되고, 식 (4.8)의  $E(n) = \alpha E(t^\beta)$ 인 관계를 활용하여  $E(Z^2)$ 에 대해 정리하면,

$$\begin{aligned} E(Z^2) &= \frac{\text{Var}(n) - \alpha E(t^\beta) + \alpha^2 E(t^\beta)^2}{\alpha^2 E(t^{2\beta})} \\ &= \frac{E(n^2) - E(n)}{E(n)^2} \cdot \frac{E(t^\beta)^2}{E(t^{2\beta})} \end{aligned} \quad (4.26)$$

이 된다. 그런데, Frailty 추가모형에서  $E(Z) = 1$ 이므로, 결국  $\text{Var}(\hat{\alpha})$ 는

$$\text{Var}(Z) = \frac{E(n^2) - E(n)}{E(n)^2} \cdot \frac{E(t^\beta)^2}{E(t^{2\beta})} - 1 \quad (4.27)$$

로 표현된다(Slimacek and Lindqvist (2016)). 물론 실제  $\text{Var}(\hat{\alpha})$ 의 추정치는 실험식인 식 (4.28)을 활용하여 추정하게 된다.

$$\widehat{\text{Var}}(Z) = \frac{\sum_{j=1}^N n_j^2 - \sum_{j=1}^N n_j}{(\sum_{j=1}^N n_j)^2} \cdot \frac{(\sum_{j=1}^N \hat{t}_j^\beta)^2}{\sum_{j=1}^N \hat{t}_j^{2\beta}} - 1 \quad (4.28)$$

## 4.2 관측 가능한 이질요소(공변량) 추가 반영

4.1절에서 제시된 1차 확장 모형에 공변량을 추가하는 방법은 Slimacek and Lindqvist (2017)의 아이디어를 채택하였다. 즉, 강도함수에  $\exp(\eta'_0 Y)$ 를 추가하는 방법이다. 이는 4.1절의 모형과 마찬가지로 기존에 확인된 모형의 특성을 그대로 이용할 수 있는 장점이 있고, 본 연구의 수행전략인 “점진적 일반화”라는 것과도 일치되는 방법이기 때문이다. 이 같은 개념에 따라 수정된 강도함수는 식 (4.29)와 같으며, 지시변수  $Y$ 와 지시변수(Indicator variable)의 영향력을 표시하는  $\eta_0$ 가 추가된다. 각 변수는 각각  $(m \times 1)$ ,  $(1 \times m)$ 행렬로 표현된다.

$$\lambda(t) = Z\alpha\beta t^{\beta-1}\exp(\eta'_0 Y) \quad (4.29)$$

### 4.2.1 공변량 추가모형의 모수추정

우선,  $\beta$ 의 추정과 관련하여, Slimacek and Lindqvist (2017)의 연구에서 공변량이 식 (4.29)의 형태로 Frailty와 함께 고려될 경우에도 Frailty만 고려한 모형 추정값  $\hat{\beta}$ 는 변하지 않음이 증명된 바 있다. 따라서 본 논문에서도 4.2절에서 추정된 식 (4.6)의  $\hat{\beta}$ 를 별도의 증명없이 공변량 추가 모형에도 그대로 사용한다.

다음으로,  $\alpha$ 와  $z_j$ 의 추정과 관련하여, 변환  $\tilde{Z} = Z\exp(\eta'_0 Y)$ 을 적용할 경우 앞서 제시된 Frailty만을 고려하는 모형과 일치한다. 이에  $\eta_0$ 만 추정된다면 별도의 추정식이 필요 없음을 알 수 있다. 따라서  $\hat{\alpha}$ 는 식 (4.13)에 위와 같은 변환을 반영하여

$$\hat{\alpha} = \frac{1}{N} \sum_{j=1}^N \frac{n_j}{t_j^{\hat{\beta}} \exp(\hat{\eta}'_0 Y)} \quad (4.30)$$

으로 산출할 수 있다. 또한, 이와 유사한 방법으로 장비별 Frailty,  $z_j$ 의 추정값은

$$\hat{z}_j = \frac{n_j}{\hat{\alpha} t_j^{\hat{\beta}} \exp(\hat{\eta}_0' Y)} \quad (4.31)$$

과 같이 산출된다.

마지막으로 공변량 산출에는 NHPP의 기본성질을 활용한다. NHPP의 누적강도함수가  $\Lambda(t)$ 일때, 이벤트 발생량의 기대값은  $E(n) = \Lambda(t)$ 를 만족한다. 다만 공변량과 Frailty를 모두 고려하는 모형에서 기대값은 공변량,  $\eta_0$ 와 Frailty,  $Z$ 를 전제로 한 값이므로,  $E(n)$ 이 아닌 조건부 기대값  $E(n|Z, Y, t)$ 을 의미하게 된다. 이 같은 사항을 전제로  $E(n)$ 과 약간의 변형식  $E\left(\frac{n}{t^{\beta} \exp(\eta_0' Y)}\right)$ 는

$$\begin{aligned} E(n) &= E\left(E(n|Z, Y, t)\right) \\ &= E\left(Z \alpha t^{\beta} \exp(\eta_0' Y)\right) \\ &= E(Z) \alpha E(t^{\beta}) E(\exp(\eta_0' Y)) \end{aligned} \quad (4.32)$$

$$\begin{aligned} E\left(\frac{n}{t^{\beta} \exp(\eta_0' Y)}\right) &= E\left(E\left(\frac{n}{t^{\beta} \exp(\eta_0' Y)} | Z, Y, t\right)\right) \\ &= E\left(\frac{E(n|Z, Y, t)}{t^{\beta} \exp(\eta_0' Y)}\right) \\ &= E\left(Z \alpha\right) = E(Z) \alpha \end{aligned} \quad (4.33)$$



으로 산출되며, 식 (4.32)의 결과를 식 (4.33)에 반영하면,

$$E(n) = E\left(\frac{n}{t^\beta \exp(\eta'_0 Y)}\right) E(t^\beta) E(\exp(\eta'_0 Y)) \quad (4.34)$$

를 얻을 수 있다. 양변을  $E(n)$ 으로 나누고 정리하면,

$$E\left(\frac{n}{t^\beta \exp(\eta'_0 Y)}\right) \frac{E(t^\beta) E(\exp(\eta'_0 Y))}{E(n)} - 1 = 0 \quad (4.35)$$

와 같은 최종적인 관계를 얻을 수 있다. 만약 식 (4.35)의 좌변에서  $\eta_0$ 를  $\eta$ 로 바꾼 식을 식 (4.36)으로 정의한다면, 구하고자 하는 공변량,  $\eta_0$ 는  $h(\eta) = 0$ 을 만족하는 해가 된다.

$$h(\eta) = E\left(\frac{n}{t^\beta \exp(\eta' Y)}\right) \frac{E(t^\beta) E(\exp(\eta' Y))}{E(n)} - 1 \quad (4.36)$$

그런데,  $E\left(\frac{n}{t^\beta \exp(\eta' Y)}\right)$ 는 식 (4.37)로 표현되므로, 이를 식 (4.36)에 반영하여  $h(\eta)$ 의 최종식 식 (4.38)을 얻는다.

$$\begin{aligned} E\left(\frac{n}{t^\beta \exp(\eta' Y)}\right) &= E\left(E\left(\frac{n}{t^\beta \exp(\eta' Y)} \middle| Z, Y, t\right)\right) \\ &= E\left(\frac{E(n|Z, Y, t)}{t^\beta \exp(\eta' Y)}\right) \\ &= E\left(\frac{Z \alpha t^\beta \exp(\eta'_0 Y)}{t^\beta \exp(\eta' Y)}\right) \\ &= E(Z) \alpha E\left(\exp\left((\eta_0 - \eta)' Y\right)\right) \end{aligned} \quad (4.37)$$

$$\begin{aligned}
h(\eta) &= \frac{E(\exp((\eta_0 - \eta)'Y))}{E(\exp(\eta_0'Y))} \frac{E(\exp(\eta'Y))}{E(\exp(\eta_0'Y))} - 1 \\
&= \frac{E(\exp((\eta_0 - \eta)'Y)) E(\exp(\eta'Y))}{E(\exp(\eta_0'Y))} - 1
\end{aligned} \tag{4.38}$$

식 (4.38)의 분자는,  $h(\eta)$ 는  $\eta = \frac{\eta_0}{2}$ 에 대해 대칭인 성질을 확인할 수 있다. 만약,  $U, V$ 가

$$U = \exp\left[\left(\frac{\eta_0}{2} - \eta\right) \frac{X}{2}\right] \tag{4.39}$$

$$V = \exp\left[\left(\frac{\eta_0}{2} + \eta\right) \frac{X}{2}\right] \tag{4.40}$$

라면,

$$UV = \exp\left[\left(\eta_0' \frac{X}{2}\right)\right] \tag{4.41}$$

이고, 널리 알려진 Cauchy-Schwarz 부등식,  $[E(UV)]^2 \leq E(U^2)E(V^2)$ 으로 부터,

$$h\left(\frac{\eta_0}{2} \pm \eta\right) > h\left(\frac{\eta_0}{2}\right) \quad \forall \eta \neq 0 \tag{4.42}$$

와 같은 관계를 얻을 수 있다. 이로부터  $h(\eta)$ 는  $\eta = \frac{\eta_0}{2}$ 일때 최소가 됨을 알 수 있다(Slimacek and Lindqvist (2017)) 이로부터, 식 (4.43)을 얻는다.

$$\eta_0 = 2 \underset{\eta}{\operatorname{argmin}} h(\eta) \tag{4.43}$$

실제 공변량  $\eta_0$ 의 추정은 식 (4.36)의 sample form인 식 (4.44)를 사용하며, 최적화와 관련된 수치 해석적 방법을 사용한다. 본 논문에서는 R library 중 “optimize”함수를 활용하였다.

$$\hat{\eta}_0 = 2 \underset{\eta}{\operatorname{argmin}} \left[ \left( \frac{1}{N} \sum_{j=1}^N \frac{n_j}{t_j^{\hat{\beta}} \exp(\eta' y_j)} \right) \frac{(\sum_{j=1}^N t_j^{\hat{\beta}})}{(\sum_{j=1}^N n_j)} \times \left( \frac{1}{N} \sum_{k=1}^N \exp(\eta' y_k) \right) \right] \quad (4.44)$$

이상의 결과에서, 공변량  $\eta_0$ 의 추정에는 NHPP의 기본 성질만을 이용하였고, 본 논문이 다루는 좌측절단 모형 역시 NHPP모형을 전제로 하므로, 위 산식은 본 논문에서도 별도의 변환 없이 사용 가능하다. 다만, 위 식에서  $n_j$ 의 실제값을 알 수 없기 때문에, 식 (4.6)과 E-M 알고리즘을 이용하여 산출된 모수를 이용한 추정값으로 대체하여 식을 활용한다. 이상의 과정은 그림 18.에 공변량 추정절차가 추가된 형태로 그림 19.와 같다.

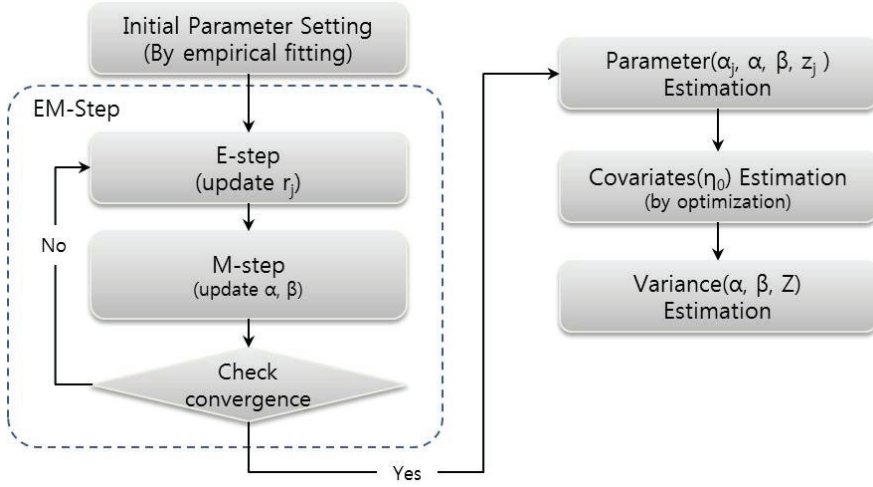


그림 19: 공변량을 추가 고려한 확장모형의 모수추정 절차

## 4.2.2 Frailty의 크기 추정

NHPP의 성질인  $E(n|Z, Y, t) = Z\alpha t^\beta \exp(\eta'_0 Y)$ ,  $\text{Var}(n|Z, Y, t) = Z\alpha t^\beta \exp(\eta'_0 Y)$ 와,  $E(Z) = 1$ 인 가정을 이용하면,  $Z$ 의 분산은 식 (4.45)로 표현할 수 있다. 세부 산식은 Appendix B.에 별도로 제시하였다.

$$\begin{aligned}
 \text{Var}(Z) &= \frac{E(n^2) - E(n)}{a^2 E(t^{2\beta}) E(\exp(2\eta'_0 Y))} - 1 \\
 &= \frac{E(n^2) - E(n)}{E(n)^2} \cdot \frac{E(t^\beta)^2 E(\exp(\eta'_0 Y))^2}{E(t^{2\beta}) E(\exp(2\eta'_0 Y))} - 1 \quad (4.45)
 \end{aligned}$$

실제  $\text{Var}(Z)$ 는 식 (4.45)의 sample식인 식 (4.46)을 사용한다.

$$\begin{aligned}
\widehat{\text{Var}}(Z) &= \frac{\frac{1}{N} \sum_{j=1}^N n_j^2 - \frac{1}{N} \sum_{j=1}^N n_j}{\left(\frac{1}{N} \sum_{j=1}^N n_j\right)^2} \times \frac{\left(\frac{1}{N} \sum_{j=1}^N t_j^{\hat{\beta}}\right)^2}{\frac{1}{N} \sum_{j=1}^N t_j^{2\hat{\beta}}} \\
&\quad \times \frac{\left(\frac{1}{N} \sum_{j=1}^N \exp(\hat{\eta}_0' Y_j)\right)^2}{\frac{1}{N} \sum_{j=1}^N \exp(2\hat{\eta}_0' Y_j)} - 1 \\
&= \frac{1}{N} \frac{\sum_{j=1}^N n_j^2 - \sum_{j=1}^N n_j}{\left(\sum_{j=1}^N n_j\right)^2} \times \frac{\left(\sum_{j=1}^N t_j^{\hat{\beta}}\right)^2}{\sum_{j=1}^N t_j^{2\hat{\beta}}} \\
&\quad \times \frac{\left(\sum_{j=1}^N \exp(\hat{\eta}_0' Y_j)\right)^2}{\sum_{j=1}^N \exp(2\hat{\eta}_0' Y_j)} - 1 \quad (4.46)
\end{aligned}$$

만약, 개별 장비들의 Frailty 크기,  $\text{Var}(z_j)$ 의 측정이 필요하다면, 식 (4.47)의 관계를 이용할 수도 있다.

$$\begin{aligned}
\text{Var}\left(\frac{n}{\alpha t^{\beta} \exp(\eta_0' Y)}\right) &= E\left[\text{Var}\left(\frac{n}{\alpha t^{\beta} \exp(\eta_0' Y)} \middle| Z, Y, t\right)\right] \\
&\quad + \text{Var}\left[E\left(\frac{n}{\alpha t^{\beta} \exp(\eta_0' Y)} \middle| Z, Y, t\right)\right] \\
&= E\left[\frac{Z}{\alpha t^{\beta} \exp(\eta_0' Y)}\right] + \text{Var}(Z) \\
&= \frac{E(Z)E(t^{-\beta})E(\exp(\eta_0' Y))}{\alpha} + \text{Var}(Z) \quad (4.47)
\end{aligned}$$

이후 실제  $\widehat{\text{Var}}(z_j)$ 의 추정에는 sample식인 식 (4.48)을 사용한다.

$$\begin{aligned}
\widehat{\text{Var}}(z_j) &= \widehat{\text{Var}}(Z) + \frac{\left(\frac{1}{N} \sum_{j=1}^N\right) \left(\frac{1}{N} \sum_{j=1}^N \exp(\hat{\eta}_0' Y_j)\right)}{\hat{\alpha}} \\
&= \widehat{\text{Var}}(Z) + \frac{\left(\sum_{j=1}^N\right) \left(\sum_{j=1}^N \exp(\hat{\eta}_0' Y_j)\right)}{N^2 \hat{\alpha}} \quad (4.48)
\end{aligned}$$

### 4.3 확장 모형의 성능 확인

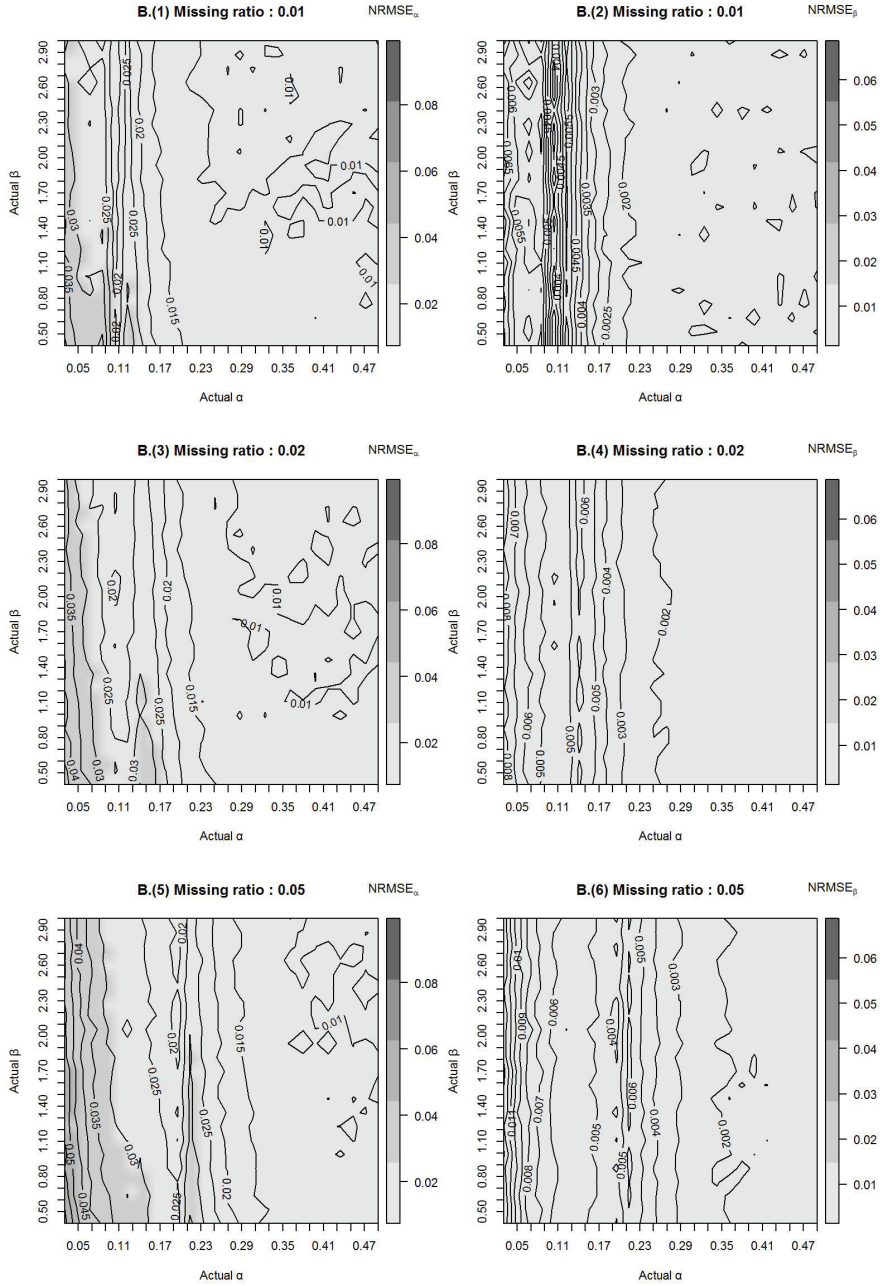
기본 모형의 민감도 확인을 위해 수행된 모의실험 조건에 Frailty와 공변량요소를 추가 후 신규 모의실험을 수행하여 제안된 모형의 특성을 확인하였고, 동일한 자료를 기본모형으로 분석한 결과와 비교하여 확장 모형의 효용성을 확인하였다. 모의실험은 표 2.의 실험조건과 유사한 방법으로 수행되었다. 우선 Frailty만 고려한 모형은 Slimacek and Lindqvist (2016)에서 사용된 방법인  $E(Z) = 1$ ,  $\text{Var}(Z) = \delta = 0.1, 0.3, 0.5, 0.7$ 인 감마 분포를 각 실험주건에 추가하여 각 경우를 358,800번씩 수행하였다. 또한, 추가로  $\delta = 0.1$ 인 경우의 모의실험 자료를 기본모형에 추가로 적용하여 해당경우의 계산 효율을 확인해 보았다. 다음으로, 공변량을 추가 고려한 모형은 편의상  $\eta = 0.5$ 인 1차원 공변량만을 사용하였고, 역시  $\delta = 0.1$ 인 경우에 대해 공변량을 고려한 모형과 기본모형에 의한 결과를 비교해 보았다. 다만, 공변량 추가모형의 경우 공변량의 유무에 따라 sample개수에 영향을 받을 수도 있기 때문에, 각 실험조건의 반복수를 2배로 하여 200회씩 반복수행하였다. 이 때, sample별  $Y$ 의 유무는 각 실험별로 난수발생을 통해 적용하였다. 결과적으로 기본모형의 성능확인에 사용된 sample의 2배인 717,600회의 실험이 실시되었다.

#### 4.3.1 모형의 민감도와 효용성

##### 4.3.1.1 Frailty만 반영된 모형

Frailty만 고려한 네 가지 모의실험 결과를 2차원 등고선 형태의 그래프로 나타낸 것은 그림 20. ~ 그림 27.과 같다. Frailty 수준별 결과로부터, 제안된 모형은 Frailty의 수준에 관계없이 매우 유사한 그래프를 확인할 수 있다. 이는 제안한 모형이 Frailty의 수준과 무관하게 유사한 추정오류를

나타냄을 의미한다. 즉, Frailty의 크기에 관계없이 항상 일관된 모수추정 능력을 나타낼 수 있다는 사실을 보여준다. 물론 앞서 제시된 기본모형의 경우와 동일하게 좌측절단 비율이 증가될 수록 추정오차의 크기는 증가되고,  $\beta$ 의 오차가  $\alpha$ 에 비해 작은 것을 알 수 있다. 반면, 기본모형에서 나타났던 특정 모수 조합에 오차가 집중적으로 발생하는 현상은 발견되지 않는다. 다만, 모든 조합에서 실제  $\alpha$ 가 작을 수록 오차는 커지는 경향을 나타내고 있다.





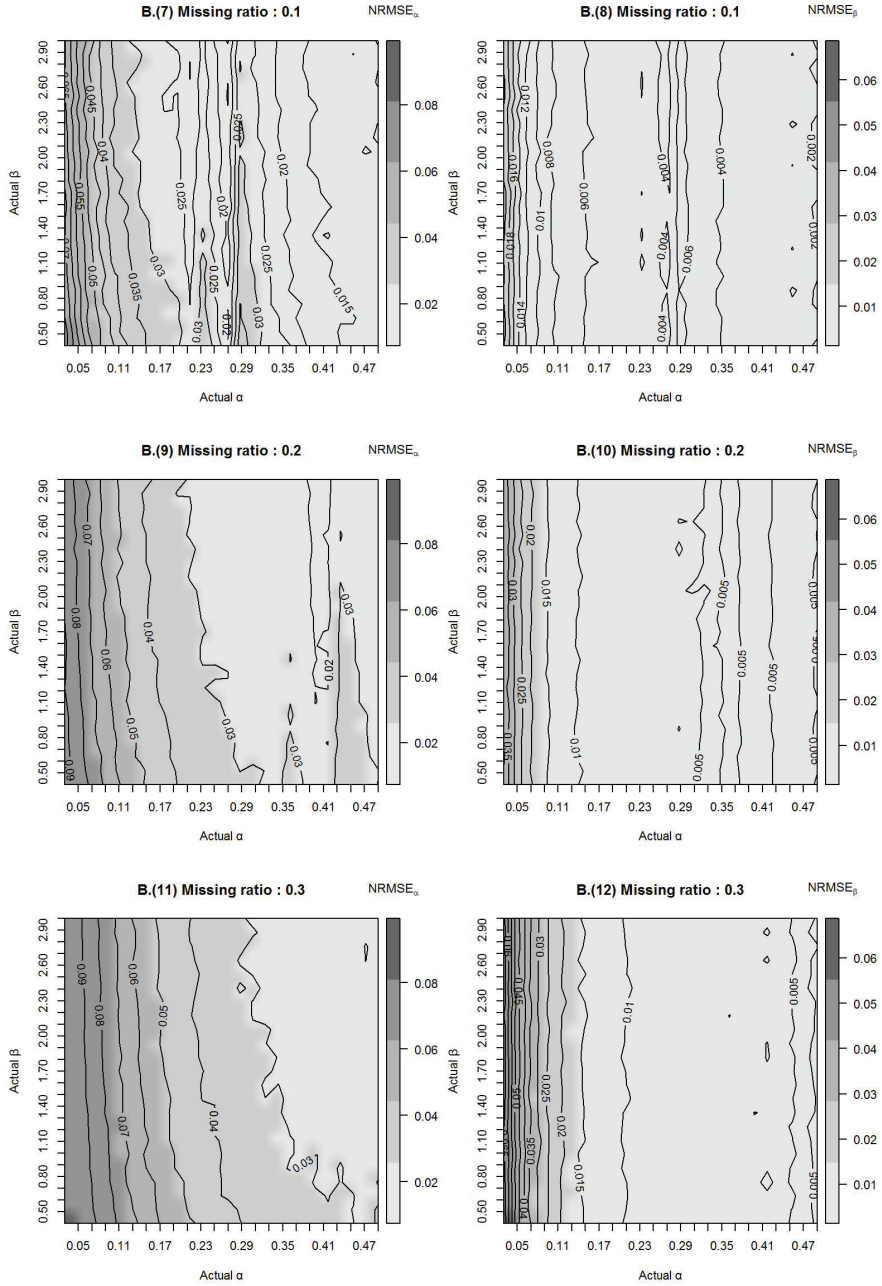


그림 21: 모의실험 결과  $NRMSE(\delta = 0.1, MR = 0.1, 0.2, 0.3)$

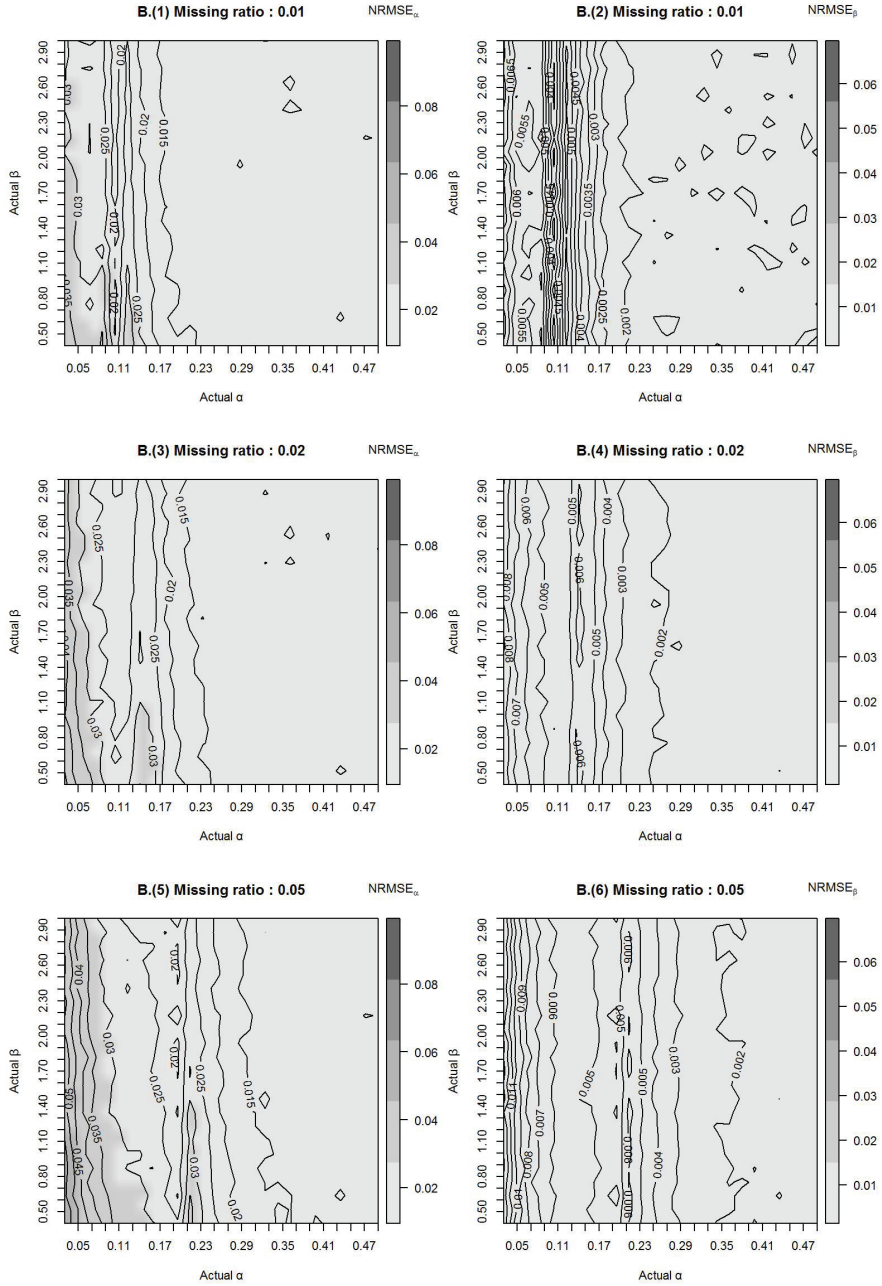


그림 22: 모의실험 결과  $NRMSE(\delta = 0.3, MR = 0.01, 0.02, 0.05)$

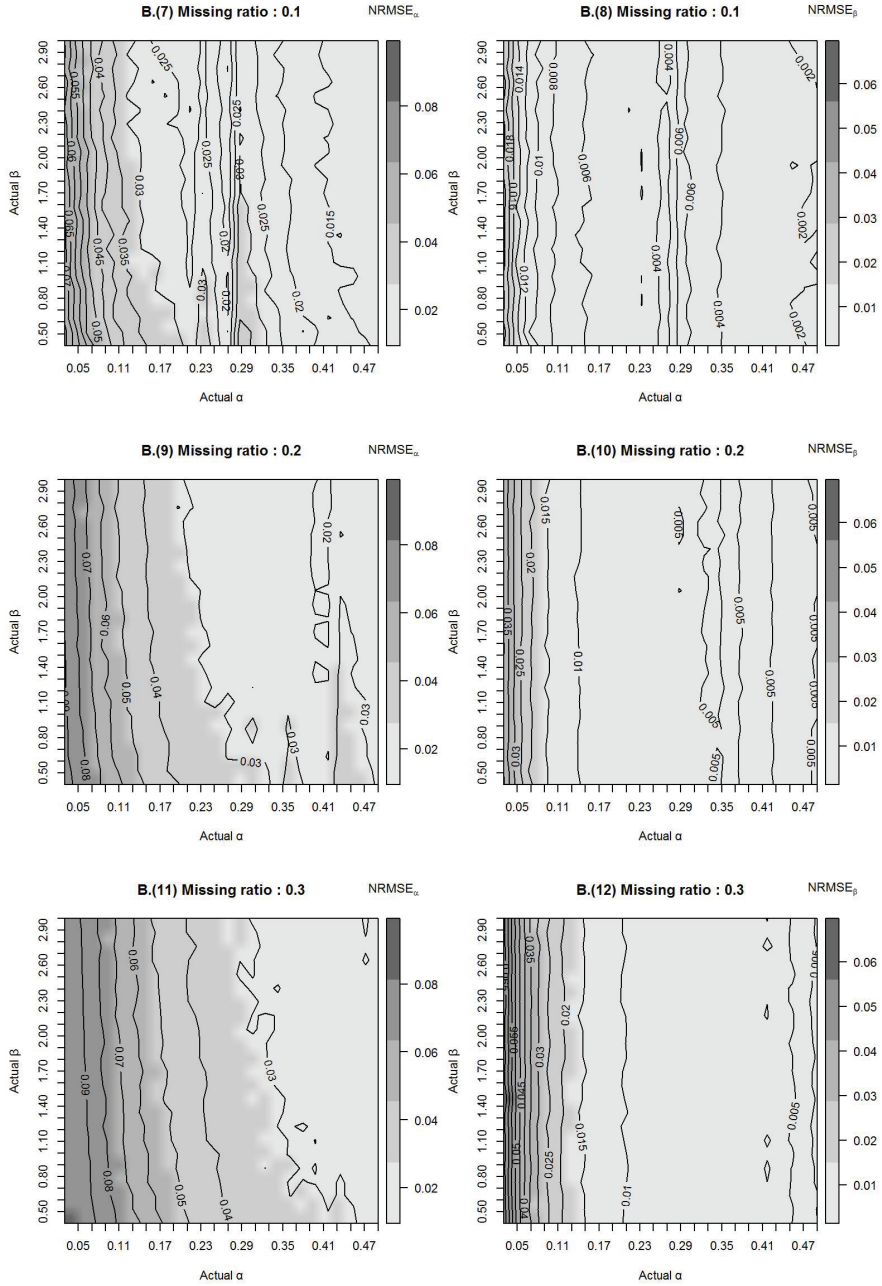


그림 23: 모의실험 결과  $NRMSE(\delta = 0.3, MR = 0.1, 0.2, 0.3)$

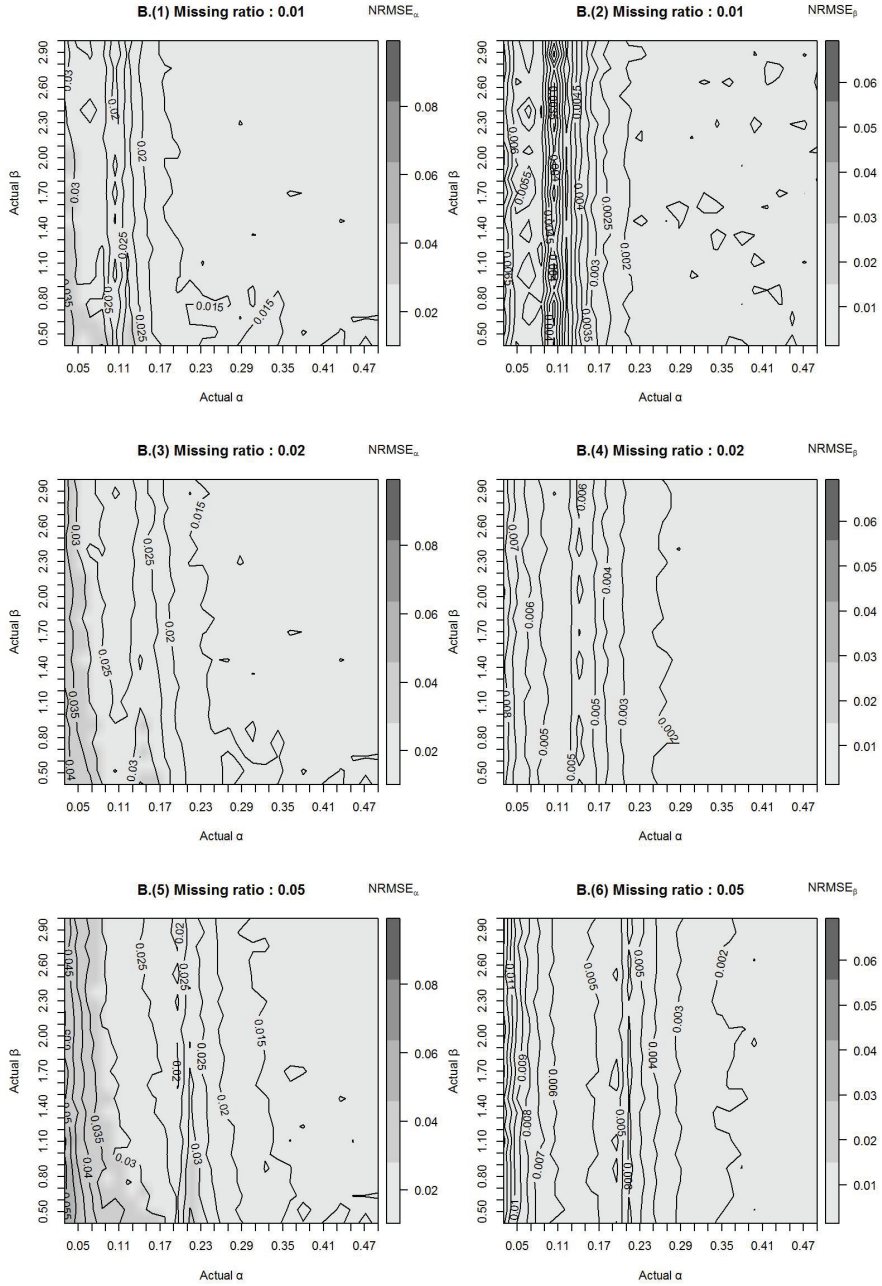


그림 24: 모의실험 결과  $\text{NRMSE}(\delta = 0.5, MR = 0.01, 0.02, 0.05)$

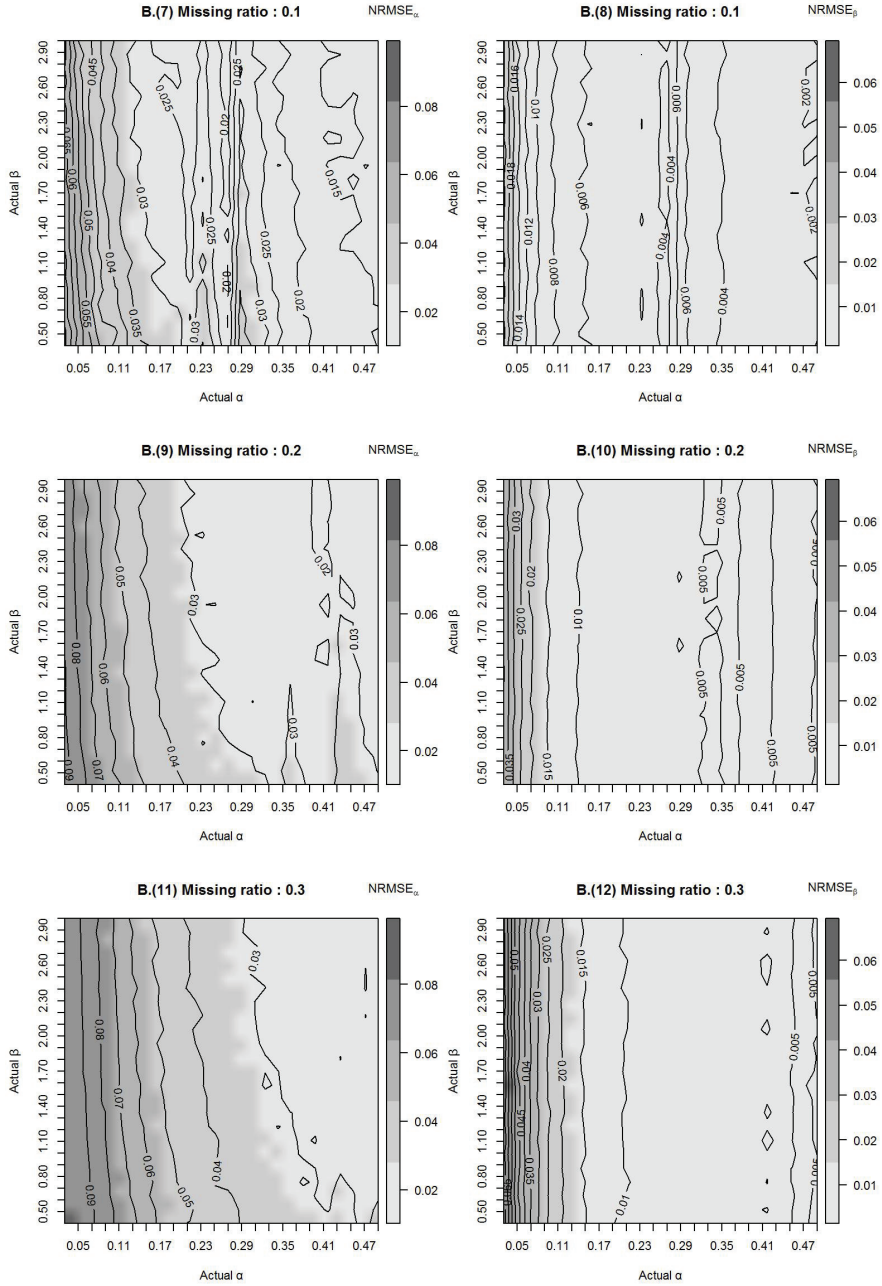


그림 25: 모의실험결과  $NRMSE(\delta = 0.5, MR = 0.1, 0.2, 0.3)$

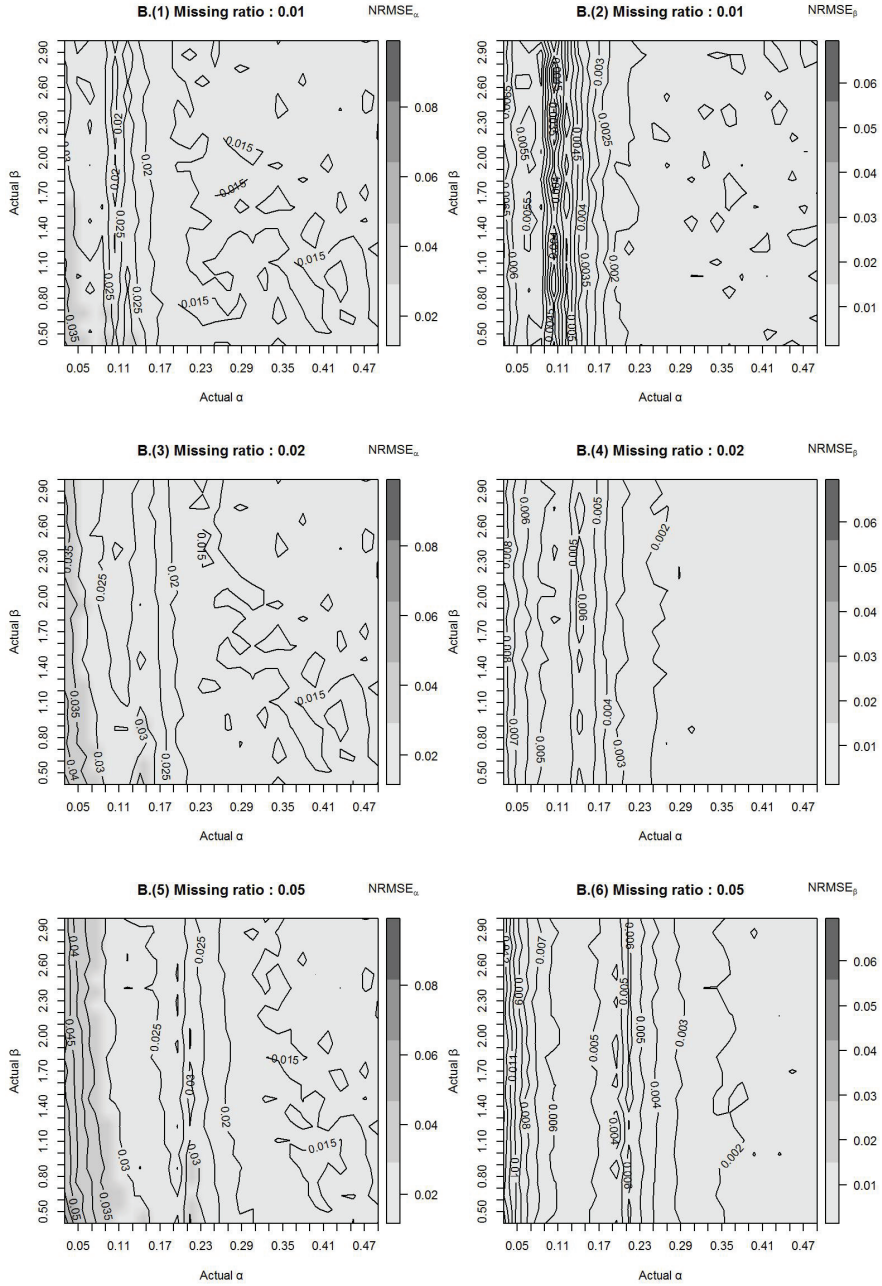


그림 26: 모의실험 결과  $NRMSE(\delta = 0.7, MR = 0.01, 0.02, 0.05)$



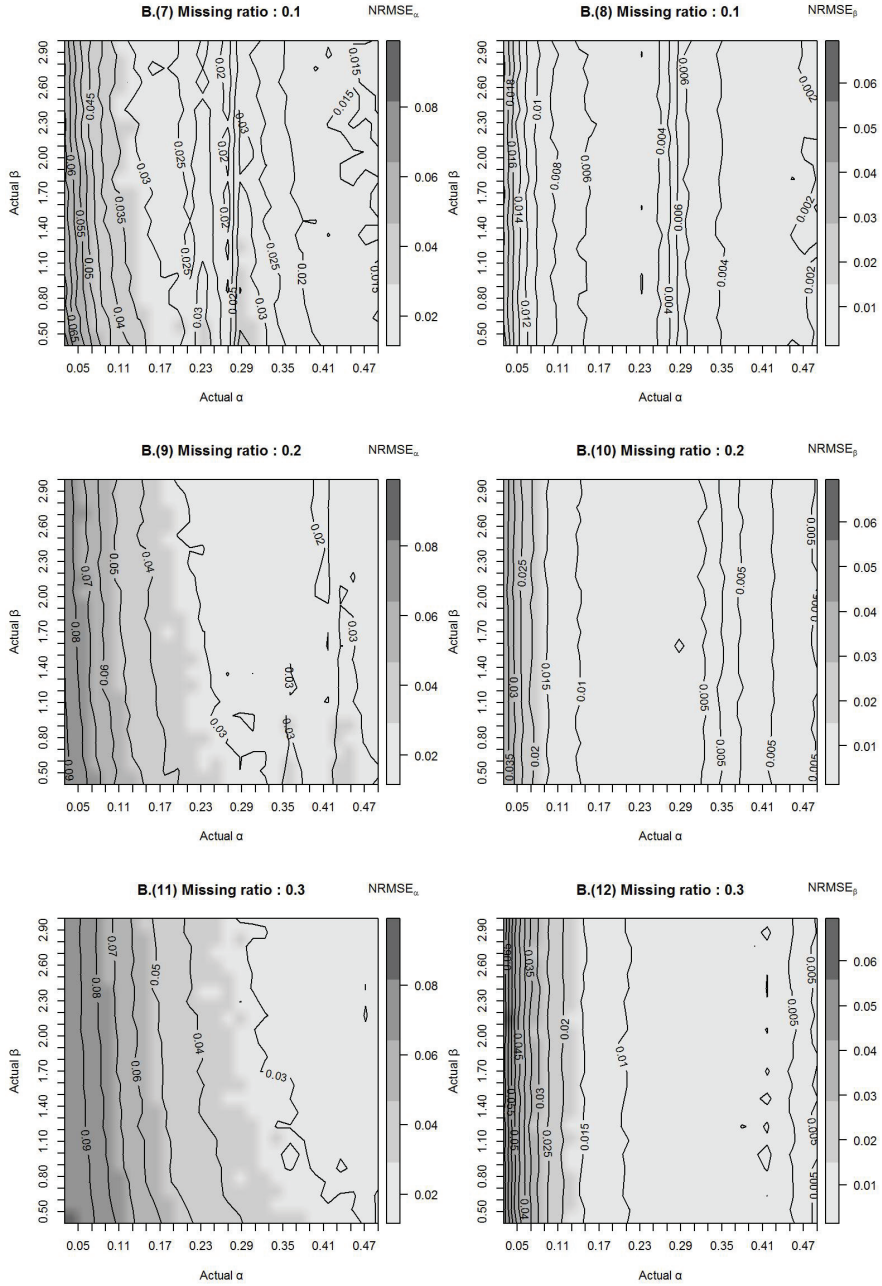


그림 27: 모의실험 결과  $NRMSE(\delta = 0.7, MR = 0.1, 0.2, 0.3)$

Frailty 수준별로 손실비율에 따른 추정효율을 비교한 결과는, 표 6., 표 7.과 같으며, 앞서 언급한 바와 같이, Frailty의 수준보다는 절단 비율에 더 큰 영향을 받고 있음을 쉽게 확인할 수 있다.

표 6: Frailty 추가모형의 모의실험 결과( $\overline{NRMSE}_\alpha$ )

Frailty 수준		0.1	0.3	0.5	0.7
손실수준	0.01	0.0152	0.0164	0.0174	0.0183
	0.02	0.0171	0.0180	0.0190	0.0198
	0.05	0.0215	0.0221	0.0228	0.0233
	0.1	0.0278	0.0280	0.0284	0.0286
	0.2	0.0384	0.0383	0.0383	0.0382
	0.3	0.0468	0.0466	0.0465	0.0463
평균		0.0278	0.0282	0.0287	0.0291

표 7: Frailty 추가모형의 모의실험 결과( $\overline{NRMSE}_\beta$ )

Frailty 수준		0.1	0.3	0.5	0.7
손실수준	0.01	0.0028	0.0028	0.0028	0.0028
	0.02	0.0032	0.0032	0.0032	0.0032
	0.05	0.0042	0.0042	0.0042	0.0042
	0.1	0.0059	0.0059	0.0059	0.0059
	0.2	0.0096	0.0096	0.0096	0.0095
	0.3	0.0144	0.0144	0.0144	0.0144
평균		0.0067	0.0067	0.0067	0.0067

손실수준에 따라 오차는 점차 증가되나, 오차 증가 속도는 기본모형과 비교할 때 급격히 증가되지 않는 것을 알 수 있으며,  $\alpha$ 의 경우 평균 약 3%,  $\beta$ 의 경우 평균 1%정도의 오차를 보이고 있다. 이는 추정효율 면에서 매우 중요한 의미를 나타낸다 할 수 있다. 즉, 기본모형의 경우 손실비율 10% 이상일 경우 정량적 분석의 실효성이 낮을 수 있으나, 확장모형의 경우 손실비율 10%이상에서도 일정수준 유의미한 결과를 얻을 수 있기



때문이다. 따라서, 좌측절단이 발생된 자료를 분석 할 때, Frailty의 존재 유무에 관계없이 확장모형을 우선 적용하는 것이 유리하다는 결론을 얻을 수 있다.

위 결과로부터, 기본모형 적용결과와의 비교는 Frailty의 크기에 따라 비교할 필요는 없음을 알 수 있다. 이에 따라 이후 분석되는 공변량 추가모형과 기본모형과의 비교는 실험에 사용했던 조건 중  $\delta = 0.1$ 인 경우만을 고려한다.

확장모형의 효용을 확인하기 위한 가장 간단한 방법은 Frailty가 없는 자료를 확장모형으로 분석 할 경우에도 유사한 결과를 얻을 수 있는지 여부를 확인하는 것이다. 만약, Frailty가 전혀 없는 자료를 확장모형으로 분석하였을 때, Frailty가 존재하는 것으로 분석되거나, 모수추정값의 차이가 실제와 큰 차이를 나타낸다면 확장모형의 의의가 없어지기 때문이다. 이를 확인하기 위해 3.3.2절의 신뢰구간의 추정에 사용한 분석자료를 두 가지 모형으로 분석하여 표 8.로 정리하였다.

표 8: 기본모형과 확장모형의 모수추정 결과 비교

구분		$\alpha$	$\beta$	$\delta$
실제값		0.1	1.2	0.0
기본 모형	추정값	0.1054	1.1884	-
	오차 (%)	5.4	0.97	-
	95% 신뢰구간	[0.0842, 0.1266]	[1.1534, 1.2234]	-
확장 모형	추정값	0.0935	1.2259	0.0192
	오차 (%)	6.5	2.2	1.9
	95% 신뢰구간	[0.0737, 0.1133]	[1.1896, 1.2623]	-

표 8.로부터 확장모형에 의한 추정결과는 실제로 없는 것으로 추정되어야 하는  $\delta$ 의 크기를 1.9%가량 있는 것으로 추정한 면이 없지 않으나,

원하는  $\alpha, \beta$ 의 추정값은 두 모형 모두 유사함을 확인할 수 있다. 즉, Frailty의 유무에 관계없이 확장모형의 사용이 유리하다는 결론을 얻는다. 다음으로, 실제 frailty가 있는 자료를 기본모형으로 분석한 경우 어느정도의 오차를 야기하는가를 분석해 보았다. 이 결과는 공변량이 추가된 모형의 결과와 함께 4.3.2절에 제시한다.

#### 4.3.1.2 두 이질요소를 모두 반영한 모형

공변량 추가모형의 경우 기본적인  $\alpha, \beta$ 추정식은 Frailty 고려모형과 동일하므로, 공변량 추가모형에 대한 Power Law 핵심모수의 추정과 관련된 민감도 분석은 별도로 수행할 필요가 없으며, 앞서 제시된 그림 20. ~ 그림 27.로 대체 가능하다. 다만, 자료손실 비율의 증기에 따라 공변량 추정값의 오차도 늘어날 것이므로, 이에 대한 결과만을 확인하였다. 앞서 제시된 여섯 가지 자료손실 비율별로 모의실험 결과의  $\eta_0$  추정결과를 표 9.로 정리하였다.

표 9: 공변량 확장모형의 손실비율 별 공변량,  $\eta_0$  추정결과

손실수준(%)	$\hat{\eta}_0$	표준편차( $\times 10^{-6}$ )
0.01	0.5000	1.6957
0.02	0.5000	1.6983
0.05	0.5000	1.7016
0.10	0.5000	1.7012
0.20	0.5000	1.6991
0.30	0.5000	1.6992

위 결과로부터, 공변량 추가모형 역시 손실 수준과 무관하게 균일한 추정력을 나타내고 있으며, 결과 또한 최초 모의실험에 사용된 값을 매우 정확하게 추정하고 있음을 알 수 있다. 이상의 결과로부터 공변량 추가모형 역시 다양한 자료손실 상황에 충분히 사용 가능함을 확인하였다.

## 4.3.2 기본모형과의 성능 비교

4.3.1.1절에서 기술한 바와 같이, 확장모형은 frailty의 크기에 관계 없이 유사한 추정력을 보이고 있다. 따라서,  $\delta = 0.1$ 인 Frailty가 추가된 자료를 기본모형과 Frailty만 고려한 확장모형으로 분석하여 결과를 비교하였다. 앞서 모의실험에 사용했던 실험조건에서 상대적으로 가장 작은  $\delta$  크기를 고려한 이유는 몇 가지 경우에 대해 모의실험을 통해 확인한 결과,  $\delta$ 가 커질 경우, 기본모형으로 추정한 결과의 오차가 과도하게 커짐을 확인했기 때문이며,  $\delta = 0.1$ 인 경우만 하더라도 확장모형에 비해 상당히 큰 오차를 나타내는 것을 확인할 수 있었기 때문이다. 기본모형과의 성능비교 목적이 확장모형이 기본모형에 비해 추정결과의 의미있는 개선효과를 줄 수 있는지 여부를 확인하는 것이므로,  $\delta = 0.1$ 인 경우에도 충분한 의미를 가질 수 있을 것으로 판단하였다. 결과는 그림 28.과 같다.

그림 28.에서, 손실비율이 작을 경우라 하더라도 확장모형이 기본모형에 비해 추정결과가 좀 더 참값에 가까운 현상을 나타내고 있다. 특히 10%이상의 손실이 발생할 경우 확장모형의 효율이 좀 더 뚜렷이 나타나며, 기본모형의  $\hat{\alpha}$  계산결과는 Frailty가 명백하게 존재하는 자료에 대해서는 매우 큰(약 2배 정도) 오차가 발생함을 알 수 있다. 이를 통해 Frailty가 존재할 가능성이 큰 경우 기본모형을 통한 정성적 분석 결과 산출은 오류 가능성이 대단히 높기 때문에 사용을 지양해야 하는것으로 판단된다. 이와 달리,  $\hat{\beta}$ 의 경우, 오히려 기본모형의 추정결과가 좀 더 나은 결과를 보이고 있다. 그러나 모든 경우에서 기본모형에 비해 확장모형의 분석 결과가 손실비율에 관계없이 균일한 경향성을 나타내며,  $\hat{\alpha}$ 와  $\hat{\beta}$ 의 오차 크기도 각 경우별로 차이가 크지 않은 점 등을 고려할 때, 확장모형의 활용도가 좀 더 나은 것으로 판단된다.

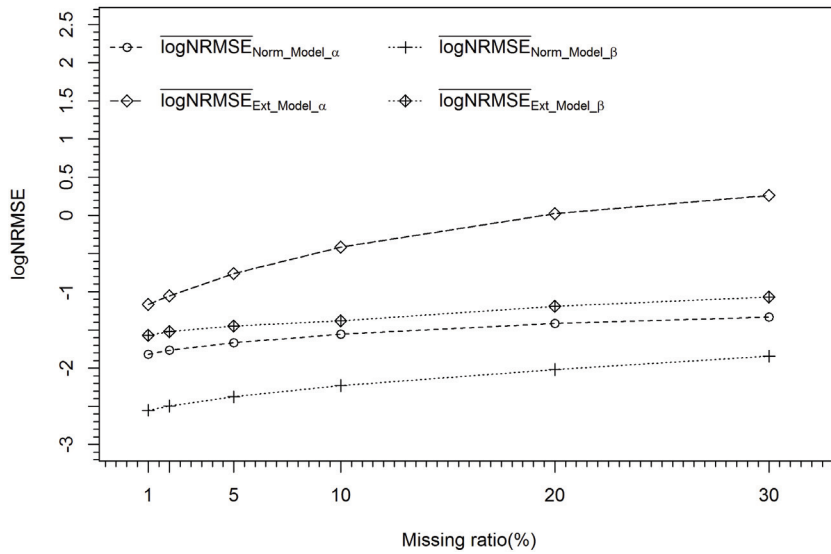


그림 28: 기본모형과 확장모형의 NRMSE 비교( $\delta = 0.1$ )

이상의 결과를 통해 Frailty가 존재하는 경우, 확장모형은 매우 유의미한 추정력을 보유하고 있으며, 기본모형에 비해 오차와 추정력의 균일성을 보장할 수 있다고 결론지을 수 있다.

## 제 5 장

# 실 자료를 활용한 실증분석

사례 분석에 사용될 대상은 '00년 초 배치되어 10년 가량 운영중인 육군의 전차이다. 분석 대상자료는 “수리부속을 필요로 하는 정비가 발생한 시점” 자료로 본 사례연구에서는 이를 “고장”으로 정의한다. 수리부속교환 실적은 직접적인 고장자료는 아니나, 대부분 고장발생에 의해 발생되므로, 통상 고장 실적으로 간주하여 분석에 사용된다. 분석에 사용된 자료는 부록 C.에 추가하였다. 분석 대상자료 중 ID 1번부터 10번까지는 좌측절단이 발생되지 않은 장비(Case I.)의 자료이며, 11번부터 56번까지는 좌측절단이 발생된 장비(Case II.)의 자료이다. 모의분석은 실제 모수를 알 수 있었고, 모의분석에 사용된 자료가 Power Law과정을 충분히 만족하고 있음을 전제로 하였다. 이와 비교하여, 실제 사례는 자료가 Power Law과정을 만족하는지 알 수 없고, 또한, 실제 모수가 무엇인지도 알 수 없다. 따라서 사례분석에서는 좌측절단이 발생하지 않은 Case I.의 자료 분석 결과를 Case II.의 비교대상으로 하여 분석을 진행하였다.

## 5.1 분석에 사용된 자료의 특성

표 10.은 부록 C.에 제시된 자료를 표 1. 형태로 정리한 것이며, 전반적인 자료의 형태와 좌측절단 정도, 장비별로 기록된 고장수 등을 나열하고 있다. Case II.에 해당하는 장비 들은 3장에서 언급된 유효한 자료손실 범위를 고려하여, 육군으로부터 확보한 전체 K계열 전차 고장자료 중에

좌측절단이 10% 이하로 발생된 장비의 자료들이다.

표 10: K계열 전차의 아전 수집자료

장비 ID	$s_j$	$t_j$	$n_j - r_j + 1$	장비 ID	$s_j$	$t_j$	$n_j - r_j + 1$
1	0.0	213.2	12	29	29.0	374.6	25
2	0.0	225.3	17	30	33.0	417.3	43
3	0.0	340.2	16	31	33.0	404.8	31
4	0.0	241.8	14	32	31.0	378.0	30
5	0.0	303.9	23	33	33.0	414.3	30
6	0.0	224.2	17	34	31.0	413.3	37
7	0.0	270.8	28	35	34.0	427.3	42
8	0.0	244.0	22	36	30.0	440.7	29
9	0.0	298.2	29	37	33.0	452.8	27
10	0.0	274.0	16	38	30.0	423.8	22
11	77.3	955.8	75	39	34.0	341.7	27
12	54.5	959.3	93	40	34.6	385.4	17
13	32.0	382.5	76	41	33.3	409.0	19
14	41.0	423.6	78	42	31.0	394.5	14
15	41.0	410.6	74	43	33.0	385.5	26
16	40.5	441.9	52	44	30.0	325.5	18
17	38.0	403.2	79	45	29.7	317.5	17
18	34.0	399.1	79	46	31.7	376.2	23
19	29.0	448.5	35	47	33.0	360.7	21
20	34.0	503.0	23	48	30.0	356.0	34
21	29.0	527.0	27	49	21.0	294.5	24
22	31.0	448.0	29	50	27.0	300.0	30
23	32.0	461.0	25	51	28.0	299.0	30
24	29.0	498.0	25	52	25.0	352.5	26
25	29.0	452.3	28	53	26.0	304.0	17
26	30.0	492.0	29	54	23.0	322.0	23
27	30.0	481.0	36	55	25.5	308.5	17
28	32.0	461.7	22	56	23.0	303.5	24

모형의 적용에 앞서, 분석 대상 자료가 과연 단조 강도함수를 갖는 NHPP 인지 여부를 확인하기 위한 간단한 테스트를 수행하였다. 이를 확인하는 방법에는 여러가지가 있으나 Laplace test와 Crow test(Kvaløy and

Lindqvist (1998); Bae et al. (2013); Bain et al. (1985))가 주로 사용되고 있음을 확인할 수 있어 두 가지 테스트를 본 논문에서도 실시하였다.

Laplace test는 아래와 같은 가설을 검정하는 방법으로,

$$H_0 : \text{HPP}$$

$$H_1 : \text{NHPP with monotonic intensity function}$$

$(a_j, b_j]$ 구간에서 발생한  $n_j$ 개의 고장시점  $x_{ji}$ 에 대해

$$L_C = \frac{\sum_{j=1}^N \sum_{i=1}^{n_j} x_{ji} - \sum_{j=1}^N \frac{1}{2} n_j (b_j + a_j)}{\sqrt{\frac{1}{12} \sum_{j=1}^N n_j (b_j - a_j)^2}} \quad (5.1)$$

과 같은 형태의 검정 통계량을 가지며,  $L_C \sim N(0, 1^2)$ 임이 알려져 있다.

이와 유사한 검정 방법인 Crow test는

$$H_0 : \text{HPP}$$

$$H_1 : \text{NHPP with increasing trend NHPP with Power Law intensity}$$

인 가설을 검정하기 위해,

$$M_C = 2 \sum_{j=1}^N \sum_{i=1}^{n_j} \ln \left( \frac{b_j - a_j}{x_{ji} - a_j} \right) \quad (5.2)$$

형태의 검정통계량을 가지며  $M_C \sim \chi^2(2 \sum_{j=1}^N n_j)$ 임이 알려져 있는 검정 방법이다.

두 방법 모두 분석 대상자료가 단조함수 형태의 NHPP 인지 여부를 확인하기 위해 널리 활용되고 있으며, 분석대상자료에 적용한 결과는 표

11.과 같다. 이를 통해 두 그룹 모두 단조함수 형태의 강도함수를 갖는 NHPP임을 확인할 수 있었다.

표 11: 실 자료의 NHPP여부 확인 결과

		Laplace test	Crow test
분포가정		$N(0, 1^2)$	$\chi^2(df = 2 \sum_j n_j)$
Case I	검정통계량	2.57	10.38
	p-value	< 0.01	< 0.001
Case II	검정통계량	8.46	22.52
	p-value	< 0.001	< 0.001

## 5.2 기본모형 적용결과

우선 Case I.의 자료를 Power Law과정으로 분석하여 모형의 적합성 여부판단의 기준값을 확보하였다. 이후 Case II.에 3장에서 제안한 모형을 적용하여 모수추정값과 신뢰구간을 산출하였다. 결과는 표 12.로 정리된다.

표 12: 실 자료의 모수추정 결과

		$\alpha$	$\beta$
Case I	모수추정값	0.0141	1.2957
Case II	모수추정값	0.0173	1.2611
	신뢰구간(95%)	[0.0112, 0.0235]	[1.2039, 1.3183]
	신뢰구간(99%)	[0.0093, 0.0254]	[1.1859, 1.3363]

표 12.의 결과로부터 Case I.과 Case II.의 모수추정결과는 상당히 유사하다는 사실을 알 수 있다. 또한, 3장에서 제안한 방법에 따라 추정된



신뢰구간 내에 Case I.의 모수추정결과가 충분히 포함되는 것을 확인할 수 있었다. 이 같은 사실을 통해 두 그룹은 동질성을 가지고 있다고 판단할 수 있었으며, 좌측절단이 발생된 자료를 이용하여도 손실이 발생하지 않는 자료를 사용하는 것과 유사한 결과를 산출 할 수 있음을 확인할 수 있었다. 다음으로, 신뢰도 관련 지표에 관한 결과를 확인하였다. 그림 29. 는 Case II.의 자료를 이용하여 산출된 누적 고장수  $m(t)$ 의 시간에 따른 추정값과 관련 신뢰구간을 도식화 한 것이다.

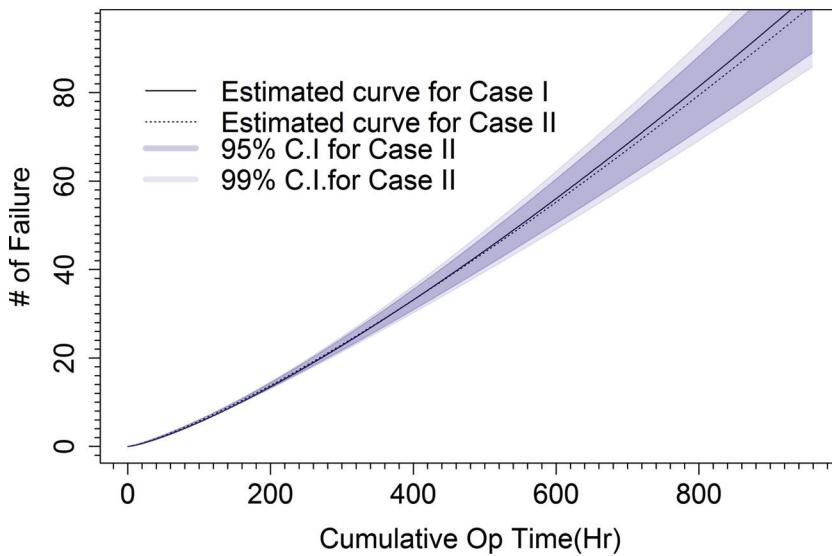


그림 29: Case I., II.의  $\widehat{m(t)}$ 와 신뢰구간 비교

표 12.의 산출결과를 이용하여 Case I.과 Case II.의  $m(t)$ 를 비교한 결과, 매우 유사한 추세를 나타내고 있음이 확인되며, 역시 Case II.의 신뢰구간 내에 Case I.의  $m(t)$  추정값이 충분히 포함됨을 알 수 있다. 추정값 외에도 추정에 사용된 실제 자료들을 동일한 그림에 표시해 본 그림을 그

림 30.으로 제시해 보았으며, 실 자료 역시 추정된 신뢰구간 내에 대부분 포함되고 있음을 확인할 수 있다.

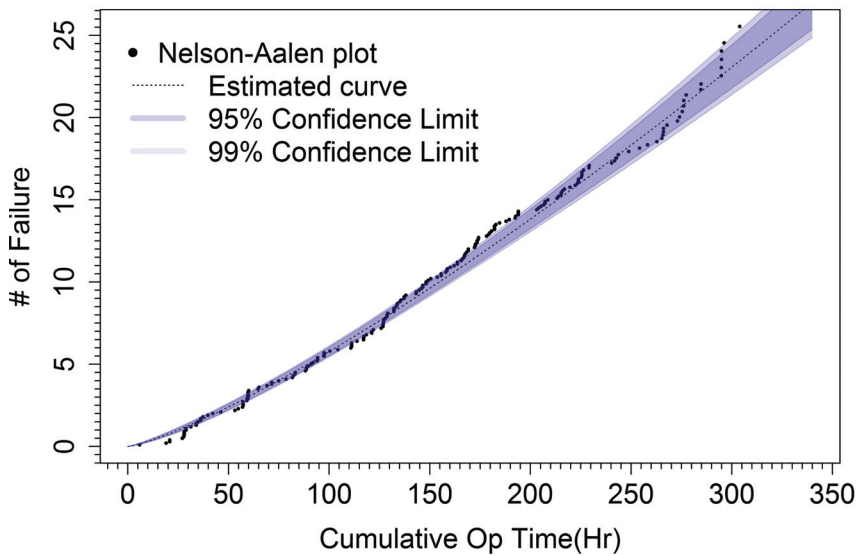


그림 30: Case I. 자료와 Case II.의  $\widehat{m(t)}$  비교

그림 30.에 나타난 점들은 개별 장비에서 발생한 고장을 각각 표시한 것은 아니다. 개별 장비의 고장을 하나씩 표시할 경우, 전반적인 비교가 곤란하기 때문에, 개별 장비 10대의 자료를 하나의 자료로 중첩하여 표현하였다. 표현에 사용된 방법은 Nelson-Aalen estimator(Nelson (1995))를 활용하였다. Nelson-Aalen estimator는 비정상자료나 절단자료(censored data)를 이용하여 비모수적 방법으로 누적위험률(cumulative hazard rate)을 구하는 방법으로 식 (5.3)을 사용한다. 생존분석이나 신뢰성 분석 분야에서 간혹 사용되는 방법으로 본 논문에서는 다수 장비에서 발생한 자료를 중첩하여 표현하기 위해 사용한다. 식 (5.3)에서  $d_i$ 는 시간 구간  $[u_i, u_j]$

에서 발생된 고장수를,  $m_i$ 는 그 때 운용되었던 장비수를 의미한다.

$$\Lambda(t) = \sum_{u_i < u_j} \frac{d_i}{m_i} \quad (5.3)$$

위 방법은 통상적으로 사용되는 방법과는 약간 상이하다. 분석 결과는 분석에 사용된 자료와 동시에 표시되는 것이 일반적인 것과 달리 본 논문에서는 분석 결과를 분석에 사용한 자료가 아닌, 분석에 사용된 것과 같은 특성값을 갖는 다른 자료와 비교하고 있다. 이는 분석에 사용된 자료는 이미  $s_j$ 이전 자료가 손실되었기 때문에, 식 (5.3)에서  $d_i$ 가 실제보다 작게 된다. 즉, 실제 관측되어야 하는 고장수보다 적은 고장수를 고려할 수밖에 없게 되어 자료가 손실된 구간에서의 고장 증가 추세가 제대로 고려될 수 없다. 이 때문에, 만약 분석에 사용된 자료를 추정된 결과 그래프와 같이 표현할 경우 그림 31.과 같이 된다. 따라서, 비록 정상적으로 추정된 결과임에도 불구하고, 항상 자료가 손실된 좌측 구간에 대해서는 추정곡선과 불일치하는 듯한 형태를 나타내게 된다. 따라서, 본 논문에서는 Case II.를 분석한 결과를 분석대상자료의 Reference라고 할 수 있는 Case I.의 자료와 비교하고 있다.

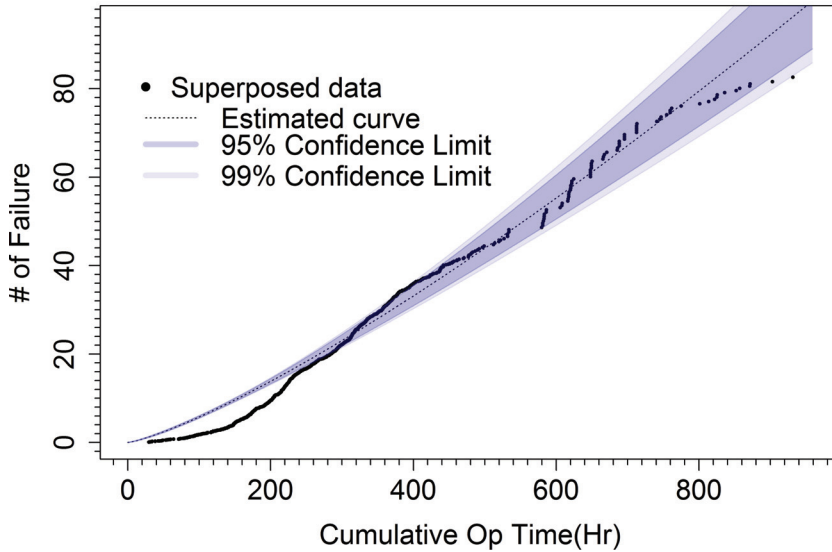


그림 31: Case II. 분석 결과와 Case II. 자료 비교

### 5.3 확장모형 적용결과

앞서 분석에 사용된 자료를 4.2절에서 제안한 확장 모형으로 분석해 보았다. Frailty만 추가된 모형은 두 요소를 모두 고려한 모형의 특별한 경우로, 별도로 비교하지 않으며, Frailty와 공변량이 모두 추가된 모형만을 사용하였다. 공변량 요소는 전차가 운용되는 지역적 특성으로 구분하였다. 장비ID 5번을 제외한 1번에서 49번까지의 장비는 A지역에서, 5번과 50번에서 56번까지 장비는 B지역에서 운용된 것으로 확인되었다. A지역에 비해 B지역의 Sample이 상대적으로 작은 이유는 A지역보다 B지역이 산지가 많기 때문에 상대적으로 전차의 운용이 제한적이기 때문이며, 대부분의 전차대대는 A지역에 위치하기 때문이다. 일반적으로 A지역의 전

차와 B지역의 전차는 상당히 다른 고장 추세를 나타낼 것으로 예상된다. 그러나 실제 운용량 분석이나 고장품목 등을 비교할 때, 아주 큰 차이를 나타내지는 않는 것으로 확인되고 있다. 만약 큰 차이가 실제로 존재하는 것이라면, 확장모형의 공변량이 유의미한 수준으로 큰 값이 산출되게 될 것이다. 결과는 표 13.과 같았다.

표 13: 확장모형의 모수추정 결과(Case II.)

구분		$\alpha$	$\beta$	$\delta$	$\eta_0$
기본모형	추정값	0.0173	1.2611	-	-
	표준편차	0.0031	0.0291	-	-
공변량 추가	추정값	0.0166	1.2593	0.0933	0.0280
	표준편차	0.0021	0.0362	-	-

표 13.에서 실 자료를 확장모형으로 분석하는 경우에도 기본모형을 적용한 결과와 큰 차이를 나타내고 있지는 않다. 이는 분석대상인 전차가 비교적 동일한 시기에 배치되었기 때문이기도 하며, 지역별로 운용 방법에 약간의 차이를 나타내기는 하지만 기본적인 운용형태는 유사하기 때문이기 때문으로 분석된다. 다만, Frailty를 의미하는  $\delta$ 가 0.0933정도 산출된 점은 전차별로 운용병의 운전습관이나 정비능력의 차이 등으로 전차의 고장특성이 완전히 동일하지는 않다는 사실을 확인시켜주고 있다. 또한 지역별 차이를 의미하는 공변량  $\eta_0$ 는 0.028로 0에 가까운 값으로 분석되고 있다. 이는 실제로 지형이나 지역적 차이에 의해 고장의 발생 시간간격이 큰 영향을 받는 것으로 생각하지만, 그 차이는 크지 않다는 것을 반증한다. 물론 세부적으로 고장이 발생하는 품목이나 위치, 발생된 고장의 임무 영향도 등은 상이할 수는 있으나, 시스템 차원에서 발생하는 누적 고장수의 차이는 큰 차이가 없음을 의미한다. 이상의 결과를 통해 분석에 사용된 K 계열 전차의 모수  $\alpha, \beta$  추정결과는 그림 32.와 같다.

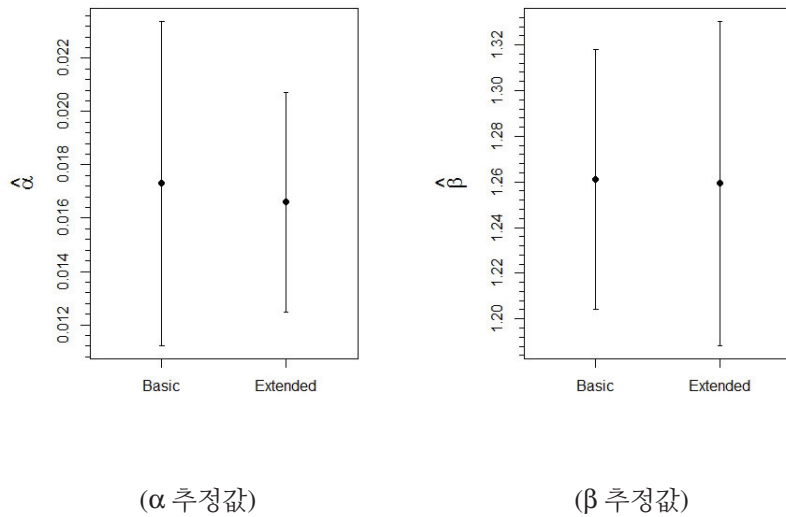


그림 32: 기본모형과 확장모형의 모수추정 결과 비교

추정값의 표준오차의 경우, 기본모형의 표준오차 중  $\alpha$ 는 기본모형이,  $\beta$ 는 확장모형이 더 큰 값을 나타내고 있으며, 4.3.2절에서 수행한 민감도 분석의 결과와 일치하는 경향임을 알 수 있다.  $\alpha$ 의 경우 확장모형에서는 기본모형에서  $\alpha$  하나로 표현되던 영향도를 비록 크지는 않으나  $Z$ 와  $\eta_0$  라는 다른 변수로 표현함으로서  $\alpha$ 의 분산을 감소시킬 수 있었기 때문으로 판단된다. 다만,  $\beta$ 의 경우 4.3.2절과 마찬가지로 오히려 확장모형에서 약간 큰 값을 나타내고 있는 바, 이는 향후 추가적인 검토가 필요할 것으로 생각된다. 다만, 두 모수 모두에서 추정값의 차이는 크지 않다. 이는 세 모형, 실제로는 확장모형 두 개의  $\alpha$ ,  $\beta$  추정치는 같으므로 두 모형의 모수추정치의 신뢰구간이 서로 겹치는 것으로부터도 쉽게 확인된다.

## 5.4 분석 결과의 활용

신뢰도 추이분석 결과를 가장 잘 활용할 수 있는 분야는 장비 관리와 정비정책 결정 분야이다. 특히 대규모의 예산이 투입되어야 하는 창정비 시점을 결정하는 분야에 유효하게 사용될 수 있다. 창정비의 업무 범위와 수행 필요성은 이미 2장에서 언급한 바와 같으며, 군에서는 가장 큰 정비정책 결정업무 중의 하나라 할 수 있다. 특히 최근에는 단순히 과거의 성능을 회복하는 수준이 아닌 무기체계의 성능개량과 연계하여 수행되는 추세로, 시점 결정에는 비용, 전략적 효용성, 전투준비태세의 영향도 등 다양한 요소가 고려되곤 한다. 다만, 다양한 요소를 고려하더라도, 발생하는 고장 간격과 이에 따르는 정비비용 증가 추이는 가장 핵심적인 요소라 할 수 있다. 통상 장기간 사용되어 우발고장만 발생하는 경우 장비의  $\beta$  추정치는 1.0 부근으로 나타난다. 만약,  $\beta > 1.0$ 인 현상이 지속적으로 발생될 경우, 장비가 노화되고 있음을 의미한다. 현재 분석 결과는  $\beta = 1.26$  부근으로 노화의 추세가 심각한 정도는 아닌 상태이다. 이는 대부분의 고장이 발생한 원인이 운용 및 정비인력의 빈번한 교체에 따른 운영미숙, 구성품의 생산 품질문제로 인해 하자보증 등이 대부분임에서도 확인된다. 실제로 이 시기 고장 내역의 대부분이 장기간 운영에 의해 발생할 수 있는 동력장치, 현수장치, 무장장치의 마모성 품목 고장이 아닌 핵심 구성품의 형상변경(Engineering Change Proposal), 예방정비 활동 중 정비 불량으로 인한 구성품 손망실, 훈련 중 부주의로 인한 장비 파손 등이 주가 되고 있다. 그러나 만약 현시점 이후에 추가로 수집된 자료를 반영하여 분석한 결과 역시 위와 같은 추세를 나타낼 경우 장비의 운용유지 측면에서 창정비를 고려해야 한다. 본 논문에서는 분석된 결과를 활용할 수 있는 고전적인 방법 두 가지를 제시한다.

우선 창정비 대상이 되는 장비가 유지해야할 신뢰도(MTBF) 목표 수준을 활용하는 방법이 있을 수 있다. MTBF 측면에서,

$$\begin{aligned}\widehat{MTBF}(480.0) &= 1/(\hat{\alpha}\hat{\beta} \times 480^{(\hat{\beta}-1)}) \\ &= 9.14Hr\end{aligned}\quad (5.4)$$

의 관계식에 따라 분석 대상 장비는 평균 운영시간인 480시간 부근에서, 매 9시간 마다 수리부속을 요하는 고장이 발생 중인 것으로 분석된다. 만약 해당 장비의 최소 요구 신뢰도 수준이 8.0시간 이상이라면, 현재는 요구 신뢰도 수준을 충분히 만족( $9.1Hr > 8.0Hr$ )하는 상태인 것으로 분석된다. 그러나 계산된  $\hat{\beta}$ 는 1.0 이상이기 때문에 시간의 흐름에 따라 신뢰도는 감소 중이고, 특정시점에 이르러서는 결국 신뢰도 수준이 요구 신뢰도 이하로 낮아질 것으로 예상된다. 본 장비의 경우 신뢰도 수준이 요구 신뢰도 수준 이하로 낮아지는 시점,  $t_0$ 는

$$\begin{aligned}MTBF(t_0) &= \frac{t_0^{1-\hat{\beta}}}{\hat{\alpha}\hat{\beta}} = 8.0 \\ \Leftrightarrow \\ t_0 &\approx 800(Hr)\end{aligned}\quad (5.5)$$

의 관계에 의해 약 800시간으로 추정된다. 따라서 군에서 요구되는 신뢰도를 만족시키기 위해서는 800시간 이전에 창정비, 특별 예방정비 활동 또는 다빈도 고장품목에 대한 주기적 교환 등 활동을 통해 고유의 신뢰도 특성을 보존하기 위해 노력할 필요가 있음을 알 수 있다. 즉, 향후 320시간이 경과되기 직전이 시점이 창정비 시점이 된다.

다음으로 단위시간당 투입되는 정비비용을 고려하는 방법이 있을 수



있다. 고장이 발생되면 반드시 정비를 수행하게 되고, 고장 빈도가 늘어남에 따라 고장정비 비용 역시 증가된다. 만약 단위시간 당 고장수가 점차 증가된다면, 정비비용 역시 증가 되므로 특정시점에는 고장이 발생하고 정비를 수행하는 것 보다 비록 고비용이 투입되지만 1회성 완전분해수리, 즉 창정비를 수행하여 고장빈도를 낮추는 것이 더 효율적일 수 있다. 이 시점을 군에서는 창정비 시점으로 하게 될 것이다. Proschan and Barlow (1967)는 위와 같은 개념을 활용하여 장비 도태시점 또는 창정비를 최소 비용으로 수행키 위한 시점을 결정하는 간단한 방법을 연구한 바 있으며, 이후 Gannon (1994); Coroporation (2009)등은 본 연구에서 적용한 수리 정책인 최소수리정책(minimal repair)을 사용하는 열화시스템에 대해 Proschan and Barlow (1967)의 방법을 적용한 연구를 수행하였다. 본 논문에서도 이와 유사한 방법으로 창정비 주기를 결정해 보았다. Proschan and Barlow (1967)가 제시한 모형은 다음과 같다.

$C_1$ 이 고장에 따른 정비비용,  $C_2$ 를 창정비 비용,  $T_{ov}$ 를 창정비 주기라 한다면, 창정비 주기 동안 발생된 총 비용,  $TC(T_{ov})$ 는

$$TC(T_{ov}) = C_1 E(N(T_{ov})) + C_2 \quad (5.6)$$

으로 표현된다. 따라서 평균 비용,  $C(T_{ov})$ 는 식 (5.7)으로, 순간 정비비용,  $IMC(T_{ov})$ 는 식 (5.8)과 같은 형태로 표현할 수 있다.

$$C(T_{ov}) = \frac{C_1 E(N(T_{ov})) + C_2}{T_{ov}} \quad (5.7)$$

$$IMC(t) = C_1 \alpha \beta t^{\beta-1} \quad (5.8)$$

만약 식 (5.7)이 식 (5.8)보다 크다면, 계속 고장정비를 수행하는 것이 유리하나, 작다면 창정비를 수행하는 것이 더 유리해진다. 따라서 두 식이 같아지는 시점의  $T_{ov}$ 가 최적 창정비 시점이 되게 된다. 즉, 아래 식 (5.9)를 만족하는 시점이 최적 창정비 시점이 된다.

$$T_{ov} = \left[ \frac{C_2}{\alpha(\beta - 1)C_1} \right]^{1/\beta} \quad (5.9)$$

군용 장비의 특성상 정확한 정비비용을 논문에 제시할 수는 없으나, 통상, 창정비 비용이 장비가의 약 60%, 연간 정비비용은 정비인력 등 부대비용을 포함하여 약 3%정도라 가정할 때, 분석 결과에 따르면 기본모형의 분석 결과 활용을 전제로, 창정비 시점은

$$\begin{aligned} T_{ov} &= \left[ \frac{60}{0.0173 \times (1.2611 - 1) \times 3} \right]^{(1/1.2611)} \\ &\approx 779.0(Hr) \end{aligned} \quad (5.10)$$

에 의해 약 779시간으로 추정된다. 즉, 현재 운용량인 480시간을 고려할 때, 향후 약 300시간 정도가 최적 창정비 시점이 된다.

이상의 결과로부터 본 논문의 사례분석에 사용된 K계열 전차의 창정비 시점은 현재로부터 약 300 ~ 320 시간 이후로 추정되며, 연간 운용량이 100시간 내외임을 고려할 때, 약 3년 후가 된다. 물론 앞서 언급한 바와 같이 최종적인 창정비 시점은 정비기간 동안 발생할 수 있는 전력 공백, 예산의 확보가능성 등 다양한 요소를 고려하여 최종 결정되게 될 것이다.

## 제 6 장

# 결론 및 추후연구

### 6.1 연구 결과

제품의 첨단화, 다양화에 따른 정비비용의 증가를 효과적으로 통제하기 위해 경험 자료의 축적과 정비업무의 과학화를 위한 노력은 다양한 분야에서 지속적으로 수행중이다. 이는 국방 분야도 크게 다르지 않다. 2009년부터 장비 운용 중 발생하는 운용, 정비, 고장 자료를 웹(Web) 기반의 데이터베이스를 통해 체계적으로 수집, 관리하고 있다. 그러나 상용 장비와 달리 대부분의 군 장비는 장기간 동안 운영되었기 때문에, 위와 같은 데이터베이스의 활용 이전에 축적된 자료 중 상당수는 유실되거나 종합(Summary)된 형태로만 남아있는 실정이다. 다시 말하면, 오래된 장비는 비록 대량의 자료가 있을 수는 있지만 좌측절단 된 자료가 대부분이다. 이러한 좌측절단 자료는 고장추세 분석에 널리 사용되는 Power Law 모형으로 분석이 불가능한 측면이 있어, 대부분의 자료가 활용이 어려워 분석에 활용되지 못하거나 활용되더라도 “관심기간 중 발생된 평균 고장 수” 정도의 간단한 통계량 산출에만 사용되고 있다. 좌측절단이 매우 심한 상황에서 잔여 자료만으로 전체 추세를 모사하는 시도는 다소 위험할 수 있다. 그러나 좌측절단이 심하지 않거나 본 연구에서 사용된 자료와 같이 기록 시작시점은 알려져 있는 등 추가적인 정보를 활용할 수 있다면, 비록 정상자료에 비해서는 효용이 낮을 수 있으나, 충분히 의미 있는 자료로 사용될 수 있을 것이라는 생각에서 본 연구는 시작되었다. 본 논문에서는

동일 형태의 다수 장비에서 발생된 좌측절단된 고장자료를 이용한 신뢰도 추세분석 모형을 제안하였다. 좌측절단에 따른 자료손실을 보상하기 위해 E-M 알고리즘을 사용하였으며, E-M 알고리즘의 단점을 보완하고자 경험적 방법을 활용하는 초기값 결정 방안도 함께 제시하였다. 제안한 신뢰도 추세분석 모형의 기반 모형으로는 유사한 분석 결과와의 비교 등 활용성을 고려하여 군용장비에서 흔히 사용되는 Power Law 모형을 사용하였다. 제안된 모형의 타당성 확인을 위해 다양한 손실비율과 모수 조합을 실험조건으로 하는 모의실험으로 손실 비율과 모수 조합별 민감도 분석을 수행하는 등 모형의 적용 가능 범위를 확인 할 수 있었다. 또한, 추정된 모수를 활용하여 MTBF 등 신뢰도 관련 지표로의 변환방법과 관련 지표의 분산을 산술식으로 제시할 수 있었다.

모형은 순수한 Power Law 형태의 강도함수만을 사용하는 “기본모형”과 이에 약간의 변형, 즉 관측 불가능한 비동질성 유발요소인 Frailty를 고려한 모형과 관측 가능한 비동질성 유발요소인 공변량을 추가한 “확장모형” 두 가지로 제안하였다. 모의실험 결과를 통해 기본모형 보다는 확장모형이 좀 더 나은 모수추정 성능과 추정결과의 균일함을 보장할 수 있음을 확인하였고, 이를 통해 간접적으로 모형의 실효성을 알 수 있었다.

기본모형의 경우 좌측절단 10% 이하의 상황에서 유의미한 모수추정결과를 나타낼 수 있음을 알 수 있었으나, 10%이상의 절단 상황에서는 Power Law 모형의 모수 중  $\alpha$ 의 모수추정치 오차가 크게 증가되어 정량적 분석에는 사용하기 다소 어려울 수 있다는 사실을 확인하였다. 물론 고장 추세를 의미하는  $\beta$  추정치의 경우 실험에 그 이상의 상황에서도 안정적으로 낮은 오차를 나타내고 있음을 알 수 있어, 신뢰도 또는 고장 추세의 변동성을 확인하는 분야에는 충분히 활용할 수 있음을 알 수 있었다.

확장 모형의 경우 기본모형에 두 가지 추가인자를 고려한 것으로, 모

형의 복잡도 증가에 따라 추정력도 높아짐을 확인할 수 있었으며, 특히 좌측절단의 정도에 관계없이 균일한 추정력을 나타내고 있음을 모의실험 결과를 통해 확인할 수 있었다. 확장모형에서 고려된 공변량, Power Law 모수 및 Frailty의 분산 추정식 역시 이론적으로 산출된 산술식의 타당성을 모의실험 결과를 통해 확인하였다.

제안된 모형은 최종적으로 실제 국내에서 사용 중인 전차에서 수집된 고장자료에 적용하여, 해당 장비의 신뢰도 추세분석 결과를 산출할 수 있었다. 이를 통해 모의실험뿐만 아니라, 실제 자료의 분석에도 충분히 사용할 수 있음을 확인할 수 있었다. 다만, 앞서 수행된 모형의 적용가능 범위와 발생될 수 있는 오차 크기 등을 고려하여 좌측절단이 10%이하로 발생된 자료만을 사용한 제한점은 있었음을 밝힌다. 마지막으로 결과의 현실적 활용 측면에서, 신뢰도 추세분석 결과를 창정비 시점 결정을 위한 분석에 사용하는 두 가지 사례를 제시하여 본 논문 결과의 실효성을 나타내고자 하였다.

이상의 결과로부터 본 연구는 사용할 수 있는 자료의 형태와 특성, 사용된 모형의 일반성과 호환성 등 두 가지 측면에서 기존 연구와 차별화된 특징을 갖는다. 우선, 자료의 형태와 특성 측면에서 기존에 다루어지지 않았던 다수 장비에서 발생한 좌측절단 자료를 사용한 최초의 연구라는 점, 동일한 장비를 사용하더라도 발생할 수 있는 자료의 비동질적 특성들을 고려한 연구라는 특징을 갖는다. 다음으로 모형의 일반성과 호환성 측면에서는 군용장비 신뢰도 추세분석에 널리 사용되는 Power Law 모형의 기본성질을 유지할 수 있는 모형이라는 점이다. 본 연구는 E-M 알고리즘을 Power Law 모형에 정형화 된 형태로 적용한 최초의 모형이다 이는 본 연구에서 제안한 모형의 분석 결과가 기존 Power Law 모형으로 분석된 결과 상호 호환성을 가질 수 있다는 의미이기도 하다. 이러한 점은 향후

본 논문의 모형을 통해 실제 야전 운용자료 또는 이와 유사한 자료를 분석하는데 있어 큰 장점으로 작용할 것으로 기대된다.

## 6.2 한계점 및 추후 연구

본 논문은 초기 고장 기록의 손실을 극복하기 위한 방안을 제시하는 연구의 출발점으로서 의의를 가진다고 할 수 있다. 다만, 분석 결과가 MTBF 등 신뢰도 지표로 쉽게 변환될 수 있어 예방정비 시점의 조정, 창정비 시점의 예측 등 실무적으로 신뢰도 추세분석 결과를 활용하는 분야에 쉽게 사용될 수 있는 장점을 갖는 반면, 아래와 같은 몇 가지 측면에서 한계점과 보완 필요성을 내제하고 있는 연구이다.

우선, 기본모형의 경우 다수 장비의 좌측절단 상황을 반영해야 했기 때문에, 산식이 단순한 형태로 정리되지 못한 단점이 있다. 이에 따라 “Newton Raphson” 등 수치 해석적 방법을 사용할 수밖에 없었다. 추정식의 난해함은 모수의 분산, 신뢰구간 추정 등에도 영향을 주어, 결국 분산의 추정 역시 결정된 산식이 아닌 수치 해석적 방법을 사용할 수밖에 없었고, 이는 최초 연구 목표인 “실무적 활용의 용이성” 측면에서 단점으로 작용될 수 있다. 물론 모형의 단순화를 위해 관측 시작과 중단 시점을 모두 같은 것으로 가정하는 등의 간단한 가정을 한다면 위의 단점은 상당부분 해소될 수도 있었을 것이다. 그러나 이 같은 추가적인 가정은 모형의 “실용성” 측면에서도 다른 단점이 될 수 있기 때문에, 더 이상의 가정을 추가하지 않는 것이 연구의 최초 방향성을 잃지 않는 것으로 판단하였다. 이에 본 연구에서는 더 이상의 추가 가정을 통한 산식의 단순화는 고려하지 않게 되었다. 다행히도 확장모형은 추가적인 인자를 고려하였음에도 불구하고, 모수와 분산 추정식 모두를 단순한 산식으로 나타낼 수 있었다. 확장모형은 기본

모형의 분석 가능범위를 포함할 수 있으므로, 기본모형의 단순화에 관한 추가연구 이전까지 보완을 위한 모형으로 활용 될 수 있을 것으로 판단 된다. 다만, 군에서 보유한 기존 모형의 분석 결과는 기본적인 Power Law 모형의 결과이므로, 결과의 호환성을 유지하기 위해서는 기본모형의 추정식 단순화를 위한 노력이 향후에 요구된다 할 수 있다. 아울러, 수치해석 방법 적용에 따른 Random 요소의 추가가 “양질의 초기값”사용에 미치는 영향과 실효성에 대한 세부 검토도 이루어 져야 할 것으로 판단된다.

다음으로, 모수의 분산과 신뢰구간 추정에 사용된 방법이 대표본을 전제로 한 근사식이라는 점은 연구결과의 “실용성”측면에서는 큰 문제가 되지 않을 수 있지만 연구의 “학술적 의미”측면에서는 단점이 될 수 있다. 물론 향후 본 연구가 대상으로 하는 자료는 실증분석에서 사용된 것 이상으로 대량의 자료이기 때문에 통계적으로 큰 문제는 아닐 가능성이 크다. 다만, 이미 단일장비를 대상으로 한 Power Law 모형의 모수에 대한 분포적 특성은 오래 전부터 알려져 있고(Handbook (1981)), 다수의 연구에서 이를 활용하고 있다는 점에서, 모수의 특성에 대한 추가적인 연구는 필요할 것으로 판단된다. 따라서 본 연구의 학문적 가치를 증대시키기 위해 해당 분야의 연구가 추가로 요구된다 할 수 있다.

무엇보다, 연구결과의 적용범위 측면에서도 정교화가 필요할 것으로 예상된다. 시뮬레이션 수행을 통한 모형의 적용범위 판단결과, 기본모형의 경우 정량적 분석 결과가 의미를 갖는 범위는 10%이하의 자료손실이 발생된 경우 정도로 확인되었다. 이는 최초 연구목표에는 충분히 도달되지 못한 것으로 여겨진다. 물론 불완전 자료를 활용하여 정상자료 수준의 결과를 도출한다는 것은 태생적으로 어려움이 예상되지만, 그럼에도 불구하고 본 연구에서 확인한 범위는 다소 협소한 것으로 확인되어 여전히 상당수의 좌측절단 자료를 활용하지 못하게 되었다. 물론 확장모형에서

는 좀 더 넓은 범위의 불완전 자료 사용이 가능할 것으로 분석되었지만 호환성 측면에서 기본모형이 더 중요성을 갖는다. 때문에 기본모형의 추정력 확보를 위한 추가연구가 시급하다. 본 연구에서 수행한 모의실험이 비록 경험적으로 대부분의 분석 영역을 포함하고 있지만, 향후 모의실험 조건을 좀 더 다양화하여 사용 가능 자료범위를 좀 더 확대할 수 있는 연구 수행은 물론 E-M 알고리즘계열의 또 다른 손실자료 보상방안에 대한 고려도 필요할 것으로 판단된다. 일반적으로 E-M 알고리즘은 손실자료의 크기가 증가될 경우 수렴 속도가 매우 느려지는 것으로 알려져 있으며, 이에 따라 본 논문에서와 같이 우도함수가 복잡하거나(기본모형) 추정해야 하는 모수가 많은 경우(확장모형) 추정속도를 보완하기 위한 방안이 요구된다. E-CM(Expectation-Conditional Maximization) 알고리즘(Meng and Rubin (1993))이 하나의 대안이 될 수 있을 것으로 판단된다. E-CM 알고리즘은 E-M 알고리즘의 M-step을 더 간단한 몇 개의 조건부 최대화(Conditional Maximization-Step)으로 대체하는 개념으로, 위와 같은 제한 상황에서 계산의 복잡도를 낮출 수 있으면서도, EM과 동등한 계산 안정성과 성능을 보이는 것으로 알려져 있다.

게다가, 이번 연구에서는 다루지 못하였으나, 군에는 Power Law 모형으로 수행한 다수의 분석 결과들이 존재한다. 이는 베이지안(Bayesian) 기법 적용을 위한 토대가 될 수 있다. 실제로 분석에 필요한 대량의 자료를 확보하지 못하는 고가의 장비나 수량이 한정된 장비의 경우 모수의 추정에 베이지안 방법을 사용하기도 한다(Hamada et al. (2008); Arab et al. (2012); Wayne and Modarres (2015)). 본 연구 역시 그와 같은 방법으로 또 다른 확장이 가능할 것으로 예상된다.

마지막으로, 현재까지 군용장비에 대해서는 활용이 미미하지만, 본 연구에서 사용한 것과는 다른 형태의 강도함수 사용도 생각해 볼 수 있다.



비록, 기존 분석 결과와의 호환성 측면에서 Power Law를 사용하고 있지만, Power Law외에도 다양한 형태의 강도함수가 존재한다. Power Law가 일반적인 추세분석 모형은 될 수 있지만 장비 특성을 보다 상세히 분석하고자 한다면 좀 더 정교화 되거나 특화된 강도함수 사용도 예상해 볼 수 있다. 앞서 II장에서 제시된 바 있는 PEXP, MPLP 등이 후보가 될 수 있으며, 본 연구와 동일한 방법으로 위 모형들도 확장 될 수 있다.

이상의 사항들은 본 연구 결과의 제한점이자 한계점일 수 있으나, 또 다른 면에서는 다른 연구의 출발점이 될 수 있다. 향후 본 연구에 사용된 가정의 완화(Relaxation), 활용 가능한 자료의 추가반영, 유사장비 분석 결과를 사용한 베이지안 기법 활용 등 연구결과를 좀 더 확장한다면, “실용성”과 “학문적 의의”를 모두 갖춘 연구가 될 수 있을 것으로 예상된다.

## 참고 문헌

- Aalen, O., Borgan, O., and Gjessing, H. (2008). *Survival and event history analysis: a process point of view*. Springer Science and Business Media.
- Anderson, R. T. and Neri, L. (2012). *Reliability-centered maintenance: management and engineering methods*. Springer Science and Business Media.
- Arab, A., Rigdon, S. E., and Basu, A. P. (2012). Bayesian inference for the piecewise exponential model for the reliability of multiple repairable systems. *Journal of Quality Technology*, 44(1):28.
- Ascher, H. and Kobbacy, K. (1995). Modelling preventive maintenance for deteriorating repairable systems. *IMA Journal of Management Mathematics*, 6(1):85–99.
- Asfaw, Z. G. and Lindqvist, B. H. (2015). Unobserved heterogeneity in the power law nonhomogeneous poisson process. *Reliability Engineering and System Safety*, 134:59–65.
- Bae, S. J., Mun, B. M., and Kim, K. Y. (2013). Change-point detection in failure intensity: A case study with repairable artillery systems. *Computers and Industrial Engineering*, 64(1):11–18.
- Bain, L. J., Engelhardt, M., and Wright, F. (1985). Tests for an increasing trend in the intensity of a poisson process: a power study. *Journal of the American Statistical Association*, 80(390):419–422.
- Bell, J. L. and Bearden, S. D. (2014). Reliability growth planning based on essential function failures. In *Reliability and Maintainability Symposium (RAMS), 2014 Annual*, pages 1–6. IEEE.
- Bhattacharjee, M., Arjas, E., Nevanlinna, R., and Pulkkinen, U. (2002). Modelling heterogeneity in nuclear power plant valve failure data.
- Black, S. E. and Rigdon, S. E. (1996). Statistical inference for a modulated power law process. *Journal of Quality Technology*, 28(1):81–90.
- Bouman, B. and Van Laar, H. (2006). Description and evaluation of the rice growth model oryza2000 under nitrogen-limited conditions. *Agricultural Systems*, 87(3):249–273.

- Brauer, D. C. and Brauer, G. D. (1987). Reliability-centered maintenance. *IEEE transactions on reliability*, 36(1):17–24.
- Broemm, W. J., Ellner, P. M., and Woodworth, W. J. (2000). Amsaa reliability growth guide. Report, ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITY ABERDEEN PROVING GROUND MD.
- Calabria, R. and Pulcini, G. (1997). Bayes inference for the modulated power law process. *Communications in statistics-theory and methods*, 26(10):2421–2438.
- Cha, J. H. and Finkelstein, M. (2014). Some notes on unobserved parameters (frailties) in reliability modeling. *Reliability Engineering and System Safety*, 123:99–103.
- Chan, A. and Cannon, P. (2002). Nonlinear forecasts of? of2: variation of model predictive accuracy over time. In *Annales Geophysicae*, volume 20, pages 1031–1038.
- Coroporation, R. (2009). Reliability growth and repairable system analysis. *Tucson: ReliaSoft Publishing*.
- Cox, D. and Lewis, P. (1966). The statistical analysis of series of events.
- Cox, D. R. (1992). *Regression models and life-tables*, pages 527–541. Springer.
- Crow, L. and Basu, A. P. (1988). Reliability growth estimation with missing data. ii. In *Reliability and Maintainability Symposium, 1988. Proceedings., Annual*, pages 248–253. IEEE.
- Crow, L. H. (1974). Tracking reliability growth. Report, ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITY ABERDEEN PROVING GROUND MD.
- Crow, L. H. (1975). Reliability analysis for complex, repairable systems. Report, ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITY ABERDEEN PROVING GROUND MD.
- Dijoux, Y., Fouladirad, M., and Nguyen, D. T. (2016). Statistical inference for imperfect maintenance models with missing data. *Reliability Engineering and System Safety*, 154:84–96.

- Duane, J. (1964). Learning curve approach to reliability monitoring. *IEEE transactions on Aerospace*, 2(2):563–566.
- Ellner, P. and Herbert, N. (2013). On some properties of a reliability growth planning model. In *Reliability and Maintainability Symposium (RAMS), 2013 Proceedings-Annual*, pages 1–6. IEEE.
- Fuqing, Y., Kumar, U., and Galar, D. (2010). Reliability prediction using support vector regression. *International Journal of System Assurance Engineering and Management*, 1(3):263.
- Gannon, J. M. (1994). *Forecasting overhaul or replacement intervals based on estimated system failure intensity*. Thesis.
- Gat, Y. L. (2014). Extending the yule process to model recurrent pipe failures in water supply networks. *Urban Water Journal*, 11(8):617–630.
- Gilardoni, G. L., de Oliveira, M. D., and Colosimo, E. A. (2013). Non-parametric estimation and bootstrap confidence intervals for the optimal maintenance time of a repairable system. *Computational Statistics and Data Analysis*, 63:113–124.
- Hamada, M. S., Wilson, A., Reese, C. S., and Martz, H. (2008). *Bayesian reliability*. Springer Science and Business Media.
- Handbook, M. (1981). Reliability growth management. *Publ. MIL-HDBK-189*.
- Ho, S. and Xie, M. (1998). The use of arima models for reliability forecasting and analysis. *Computers and industrial engineering*, 35(1-2):213–216.
- Ho, S., Xie, M., and Goh, T. (2002). A comparative study of neural network and box-jenkins arima modeling in time series prediction. *Computers and Industrial Engineering*, 42(2):371–375.
- Hong, W.-C. and Pai, P.-F. (2006). Predicting engine reliability by support vector machines. *The International Journal of Advanced Manufacturing Technology*, 28(1-2):154–161.
- Karbasian, M. and Ibrahim, Z. (2010). Estimation of parameters of the power-law-non-homogenous poisson process in the case of exact failures data.

- Kumar, D. (1995). Proportional hazards modelling of repairable systems. *Quality and Reliability Engineering International*, 11(5):361–369.
- Kvaløy, J. T. and Lindqvist, B. H. (1998). Ttt-based tests for trend in repairable systems data. *Reliability Engineering and System Safety*, 60(1):13–28.
- Lakey, M. J. (1992). *The modulated power law process*. Thesis.
- Lawless, J. F. (1987). Regression methods for poisson process data. *Journal of the American Statistical Association*, 82(399):808–815.
- Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2016). A methodology for stochastic inventory models based on a zero-adjusted birnbaum-saunders distribution. *Applied Stochastic Models in Business and Industry*, 32(1):74–89.
- Lindqvist, B. H. (2006). On the statistical modeling and analysis of repairable systems. *Statistical science*, pages 532–551.
- Little, R. J. and Rubin, D. B. (2014). *Statistical analysis with missing data*. John Wiley and Sons.
- Meng, X.-L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, 80(2):267–278.
- Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, 1(2):226–251.
- Mun, B. M., Bae, S. J., and Kvam, P. (2013). A superposed log-linear failure intensity model for repairable artillery systems. *Journal of Quality Technology*, 45(1):100.
- Muralidharan, K. (2002). Reliability inferences of modulated power-law process i. *IEEE Transactions on Reliability*, 51(1):23–26.
- Na, I. Y. and Chang, W. (2017). Multi-system reliability trend analysis model using incomplete data with application to tank maintenance. *Quality and Reliability Engineering International*, 33(8):2385–2395.
- Nelson, W. (1995). Confidence limits for recurrence data—applied to cost or number of product repairs. *Technometrics*, 37(2):147–157.
- Nelson, W. B. (2009). *Accelerated testing: statistical models, test plans, and data analysis*, volume 344. John Wiley and Sons.

- Nicolardot, B., Recous, S., and Mary, B. (2001). Simulation of c and n mineralisation during crop residue decomposition: a simple dynamic model based on the c: N ratio of the residues. *Plant and Soil*, 228(1):83–103.
- O’connor, P. D., O’Connor, P., and Kleyner, A. (2012). *Practical reliability engineering*. John Wiley and Sons.
- Proschan, F. and Barlow, R. E. (1967). Mathematical theory of reliability. *SIAM series in applied mathematics*.
- Pulcini, G. (2001). Modeling the failure data of a repairable equipment with bathtub type failure intensity. *Reliability Engineering and System Safety*, 71(2):209–218.
- Ramakumar, R. (1993). *Engineering reliability: fundamentals and applications*. Prentice Hall.
- Schervish, M. J. (2012). *Theory of statistics*. Springer Science and Business Media.
- Schott, J. R. (2016). *Matrix analysis for statistics*. John Wiley and Sons.
- Slimacek, V. and Lindqvist, B. (2013). Failure rate of wind turbines modeled by homogeneous poisson process with covariates and unobserved heterogeneity. In *Safety, Reliability and Risk Analysis: Beyond the Horizon. Proceedings of the European Safety and Reliability Conference, ESREL*.
- Slimacek, V. and Lindqvist, B. H. (2016). Nonhomogeneous poisson process with nonparametric frailty. *Reliability Engineering and System Safety*, 149:14–23.
- Slimacek, V. and Lindqvist, B. H. (2017). Nonhomogeneous poisson process with nonparametric frailty and covariates. *Reliability Engineering and System Safety*, 167:75–83.
- Sun, A., Kee, E., Yu, W., Popova, E., Grantom, R., and Richards, D. (2005). Application of crow-amsaa analysis to nuclear power plant equipment performance. In *International Conference on Nuclear Engineering ICONE-13, Beijing, China, May*, pages 16–20.
- Süli, E. and Mayers, D. F. (2003). *An introduction to numerical analysis*. Cambridge university press.

- TEAM, R. C. (2010). R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. Report, ISBN 3-900051-07-0, URL: <http://www.R-project.org>.
- Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.
- Wang, P. and Coit, D. W. (2005). Repairable systems reliability trend tests and evaluation. In *Reliability and Maintainability Symposium, 2005. Proceedings. Annual*, pages 416–421. IEEE.
- Wayne, M. and Modarres, M. (2015). A bayesian model for complex system reliability growth under arbitrary corrective actions. *IEEE Transactions on Reliability*, 64(1):206–220.
- You, M. and Meng, G. (2012). Residual life prediction of repairable systems subject to imperfect preventive maintenance using extended proportional hazards model. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 226(1):50–63.
- Yu, J., Tian, G., and Tang, M. (2008). Statistical inference and prediction for the weibull process with incomplete observations. *Computational Statistics and Data Analysis*, 52(3):1587–1603.

## 부록 A. Frailty 추가모형의 분산추정 증명

임의의 행렬  $A$ 를

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (1)$$

과 같은 네 개의 블록 행렬로 구분하고,  $A$ 의 역행렬을

$$A^{-1} = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad (2)$$

로 표현할 경우, 만약  $A_{22}$ 와  $A_{11} - A_{12}A_{22}^{-1}A_{21}$ 이 비정칙행렬(singular matrix)라면, 행렬  $B$ 의 각 블록은 식 (3)에서 식 (6)으로 표현 할 수 있음

$$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \quad (3)$$

$$B_{22} = A_{22}^{-1} + A_{22}^{-1}A_{21}B_{11}A_{12}A_{22}^{-1} \quad (4)$$

$$B_{12} = -B_{11}A_{12}A_{22}^{-1} \quad (5)$$

$$B_{21} = -A_{22}^{-1}A_{21}B_{11} \quad (6)$$

식 (4.14) ~ 식 (4.18)로 표현된 Fisher Information의  $(ij)^{th}$ 인자들을

$$I = \begin{pmatrix} S & y \\ y' & x \end{pmatrix} \quad (7)$$

과 같은 블록 행렬로 표현할 때,  $S$ 는  $N \times N$  행렬,  $y$ 는  $N \times 1$  행렬이며,  $x$



는  $1 \times 1$  행렬인 대각요소에 대한 대칭인 형태의 행렬로 표현됨. 식 (2)에서 실제로 계산이 필요한 부분은  $\widehat{\text{Var}}(\hat{\alpha}_j|Z, t)$ 를 의미하는  $B_{11}$ 의 대각성분,  $\widehat{\text{Var}}(\hat{\alpha}_i, \hat{\alpha}_j|Z, t)$ 를 의미하는  $B_{11}$ 의 대각성분 외 요소, 그리고  $\widehat{\text{Var}}(\hat{\beta}|Z, t)$ 를 의미하는  $B_{22}$  뿐임. 즉, 식 (3)과 식 (4)의 계산만이 요구되므로 만약,  $I^{-1} = B$ 라면, 우선  $B_{11}$ 은 블록 행렬의 성질  $(A + CBD)^{-1} = A^{-1} - A^{-1}C(B^{-1} + DA^{-1}C)^{-1}DA^{-1}$ 에 의해,

$$\begin{aligned}
 B_{11} &= (S - \frac{1}{x}yy')^{-1} \\
 &= S^{-1} - S^{-1} \cdot (-\frac{1}{x})[(yy')^{-1} - \frac{1}{x}S^{-1}]^{-1}S^{-1} \\
 &= S^{-1} + S^{-1}[x(yy')^{-1} - S^{-1}]^{-1}S^{-1} \\
 &= S^{-1} - S^{-1}[I - xS(yy')^{-1}]^{-1}
 \end{aligned} \tag{8}$$

로 표현됨. 그런데,  $(I + A)^{-1} = I - (I + A^{-1})^{-1}$ 이므로,  $[I - xS(yy')^{-1}]^{-1}$ 는  $I - (I - \frac{1}{x}(yy')S^{-1})^{-1}$ 이 되므로,  $B_{11} = S^{-1}[I - \frac{1}{x}(yy')S^{-1}]^{-1}$ 임을 알 수 있음. 이 때,  $c, d$ 가 column 벡터인 경우  $(A + cd')^{-1} = A^{-1} - A^{-1}cd'A^{-1}/(1 + d'A^{-1}c)$ 인 성질에 의해 결국 식 (8)은 식 (9)로 정리됨.

$$\begin{aligned}
 B_{11} &= S^{-1}[I + \frac{1}{x}yy'S^{-1}/(1 - \frac{1}{x}y'S^{-1}y)] \\
 &= S^{-1} + (S^{-1}yy'S^{-1})/(x - y'S^{-1}y) \\
 &= S^{-1} + \frac{1}{x - y'S^{-1}y}(S^{-1}y)(S^{-1}y)'
 \end{aligned} \tag{9}$$

위 결과를 이용하여 구한  $B_{22}$ 는, 식 (10)의 형태가 됨.

$$\begin{aligned}
 B_{22} &= \frac{1}{x} + \frac{1}{x} y' B_{11} y \cdot \frac{1}{x} \\
 &= \frac{1}{x^2} \left[ x + y' S^{-1} y + \frac{(y' S^{-1} y)(y' S^{-1} y)'}{x - y' S^{-1} y} \right] \\
 &= \frac{1}{x^2} \left[ \frac{x^2}{x - y' S^{-1} y} \right] \\
 &= \frac{1}{x - y' S^{-1} y}
 \end{aligned} \tag{10}$$

이상의 결과를 이용하면 우선  $\widehat{\text{Var}}(\hat{\beta}|Z, t)$ 의 경우, 식 (10)에 식 (4.14)에서 식 (4.18)을 반영하고,  $\hat{\alpha}_j = \frac{n_j}{t_j^{\hat{\beta}}}$  임을 활용하면,

$$\begin{aligned}
 y' S^{-1} y &= \sum_{j=1}^N \frac{\hat{\alpha}_j^2}{n_j} t_j^{2\hat{\beta}} (\ln t_j)^2 \\
 &= \sum_{j=1}^N \hat{\alpha}_j t_j^{\hat{\beta}} (\ln t_j)^2
 \end{aligned} \tag{11}$$

이 되므로,

$$\begin{aligned}
 \widehat{\text{Var}}(\hat{\beta}|Z, t) &= \frac{1}{x - y' S^{-1} y} \\
 &= \frac{1}{K/\hat{\beta}^2 + \sum_{j=1}^N \hat{\alpha}_j t_j^{\hat{\beta}} (\ln t_j)^2 - \sum_{j=1}^N \hat{\alpha}_j t_j^{\hat{\beta}} (\ln t_j)^2} \\
 &= \frac{1}{K/\hat{\beta}^2} \\
 &= \frac{\hat{\beta}^2}{K}
 \end{aligned} \tag{12}$$

임을 알 수 있음. 마찬가지로,  $\widehat{\text{Var}}(\hat{\alpha}_j|Z, t)$ 와  $\widehat{\text{Var}}(\hat{\alpha}_i, \hat{\alpha}_j|Z, t)$ 의 경우,  $N \times 1$

행렬인  $S^{-1}y$ 의  $j$ 번째 요소가

$$\frac{\hat{\alpha}_j^2}{n_j} t_j^{\hat{\beta}}(\ln t_j) = \hat{\alpha}_j(\ln t_j) \quad (13)$$

인 것과, 식 (2)로부터, 각각

$$\begin{aligned} \widehat{\text{Var}}(\hat{\alpha}_j|Z, t) &= \frac{\hat{\alpha}_j^2}{n_j} + \frac{\hat{\beta}^2}{K} \hat{\alpha}_j^2 (\ln t_j)^2 \\ &= \frac{\hat{\alpha}_j^2}{n_j} + \widehat{\text{Var}(\hat{\beta})} \hat{\alpha}_j^2 (\ln t_j)^2 \end{aligned} \quad (14)$$

$$\begin{aligned} \widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\alpha}_j|Z, t) &= \frac{\hat{\beta}^2}{K} \hat{\alpha}_i(\ln t_i) \hat{\alpha}_j(\ln t_j) \\ &= \widehat{\text{Var}(\hat{\beta})} \hat{\alpha}_i \hat{\alpha}_j (\ln t_i) (\ln t_j) \end{aligned} \quad (15)$$

가 됨을 확인함.

## 부록 B. 공변량 추가 모형의 식 (4.45) 증명

NHPP의 성질에 의해,

$$\begin{aligned} E(n|Z, Y, t) &= Z\alpha t^\beta \exp(\eta'_0 Y) \\ \text{Var}(n|Z, Y, t) &= Z\alpha t^\beta \exp(\eta'_0 Y) \end{aligned} \quad (16)$$

과, Frailty 고려 모형의 가정인

$$E(Z) = 1 \quad (17)$$

을 이용하면,  $\text{Var}(n)$ 는 아래와 같이 정리됨.

$$\begin{aligned} \text{Var}(n) &= E\left(\text{Var}(n|Z, Y, t)\right) + \text{Var}\left(E(n|Z, Y, t)\right) \\ &= E(Z\alpha t^\beta \exp(\eta'_0 Y)) + \text{Var}(Z\alpha t^\beta \exp(\eta'_0 Y)) \\ &= E(n) + \alpha^2 \left[ E\left((Zt^\beta \exp(\eta'_0 Y))^2\right) - E\left(Zt^\beta \exp(\eta'_0 Y)\right)^2 \right] \\ &= E(n) - E(n)^2 + \alpha^2 E\left((Zt^\beta \exp(\eta'_0 Y))^2\right) \\ &= E(n) - E(n)^2 + \alpha^2 E(Z^2) E(t^{2\beta}) E(\exp(2\eta'_0 Y)) \\ &= E(n) - E(n)^2 \\ &\quad + \alpha^2 \left[ \text{Var}(Z) + 1 \right] E(t^{2\beta}) E(\exp(2\eta'_0 Y)) \end{aligned} \quad (18)$$

위 식의 좌변을  $\text{Var}(n) = E(n^2) - E(n)^2$  으로, 우변의  $\alpha$ 를

$$\alpha = \frac{E(n)}{E(t^\beta) E(\exp(\eta'_0 Y))} \quad (19)$$

으로 대체한 후 재정리 하면,

$$E(n^2) - E(n) = \alpha^2 [\text{Var}(Z) + 1] E(t^{2\beta}) E(\exp(2\beta'_0 Y)) \quad (20)$$

을 얻으며, 다시  $\text{Var}(Z)$ 에 관하여 정리하면,

$$\begin{aligned} \text{Var}(Z) &= \frac{E(n^2) - E(n)}{a^2 E(t^{2\beta}) E(\exp(2\eta'_0 Y))} - 1 \\ &= \frac{E(n^2) - E(n)}{E(n)^2} \cdot \frac{E(t^\beta)^2 E(\exp(\eta'_0 Y))^2}{E(t^{2\beta}) E(\exp(2\eta'_0 Y))} - 1 \end{aligned} \quad (21)$$

과 같이  $Z$ 의 분산을 얻게 됨.

## 부록 C. 사례분석에 사용된 야전 고장자료

- 제시된 자료는 육군의 전차 고장시점이며, 단위는 “시간(Hr)”임
- 각 장비가 운영을 시작한 시점이 “0.0” 이며, 각 고장시점은 고장이 발생했을 때까지 누적된 엔진가동시간을 의미함

장비ID	고장시점(Hr)
1	28.87, 28.87, 64.62, 118.42, 126.78, 128.78, 157.87, 158.7, 172.53, 192.42, 205.17, 213.17
2	19, 20.75, 20.75, 54.83, 59.67, 59.67, 59.67, 60, 78.58, 88.67, 100.33, 113.83, 125.83, 145.78, 203.13, 223.8, 225.3
3	74.87, 88.2, 92.03, 93.97, 94.3, 94.3, 132.15, 137.3, 173.45, 174.2, 258.1, 262.77, 265.27, 265.77, 284.6, 284.7
4	27.08, 27.83, 57.25, 57.25, 57.25, 81.75, 83.2, 133.67, 135.67, 167.55, 172.22, 219.63, 240.3, 241.8
5	34.78, 59.12, 59.78, 59.78, 65.2, 97.45, 97.45, 97.45, 117.03, 133.2, 138.08, 187.85, 193.93, 193.93, 215.07, 215.9, 216.82, 276.15, 276.15, 277.32, 294.9, 294.9, 303.9
6	42.17, 69.17, 110.83, 111.17, 111.17, 111.17, 117.17, 117.17, 128.67, 144.67, 149.17, 150.5, 153.67, 189.42, 192.67, 208.67, 222.17
7	34, 34, 37.33, 53.17, 82.17, 91.17, 126.83, 127.17, 132.17, 143.17, 143.17, 155.67, 155.67, 172.67, 180.67, 182.17, 182.17, 182.17, 182.83, 184.67, 194, 223.33, 223.83, 225.5, 225.5, 226.33, 243.67, 267.83
8	36.17, 36.17, 57, 83, 126.67, 127, 127.83, 134, 146.5, 146.5, 158.5, 162, 165.5, 166.25, 167, 169.17, 173.5, 174, 207.25, 207.25, 229, 242.67
9	28, 28, 28, 31.33, 46.17, 121.17, 121.17, 122, 129.17, 132.17, 135.17, 137.17, 148.17, 148.17, 160.17, 163.67, 163.67, 167.17, 168.17, 169, 178.17, 214, 215.33, 229.17, 273, 275.42, 294.75, 294.75, 296.17

continued ...

...continued

장비ID	고장시점(Hr)
10	5.83, 39.67, 71.5, 71.5, 88.75, 89.67, 104.33, 179.22, 179.88, 204.03, 206.55, 248.78, 254.12, 266.03, 266.03, 274.03
11	79.88, 128.53, 143.45, 143.45, 146.03, 149.95, 150.12, 150.62, 214.7, 215.37, 217.03, 220.53, 220.87, 221.12, 222.28, 259.28, 259.28, 278.2, 278.2, 278.7, 278.7, 278.7, 278.87, 281.12, 286.03, 290.45, 293.12, 293.62, 315.95, 377.45, 390.45, 391.45, 394.78, 394.78, 394.78, 395.37, 395.87, 395.87, 396.87, 397.2, 435.7, 436.28, 438.78, 496.62, 519.7, 579.48, 581.07, 581.07, 582.07, 582.82, 583.15, 583.32, 585.9, 586.23, 605.6, 607.77, 608.27, 619.7, 649.37, 649.37, 649.37, 667.58, 670.92, 685.42, 685.92, 688.17, 688.17, 688.17, 775.03, 813.43, 821.93, 824.93, 825.43, 835.77, 902.83
12	85.22, 134.73, 143.98, 146.82, 147.82, 149.15, 149.48, 150.48, 152.68, 155.52, 165.35, 202.1, 209.23, 210.32, 211.52, 211.85, 213.77, 214.77, 215.35, 216.02, 216.77, 217.85, 254.85, 278.52, 279.02, 279.02, 292.35, 304.18, 306.77, 310.43, 312.43, 313.43, 316.6, 317.43, 318.18, 318.68, 339.18, 339.18, 421.68, 423.18, 436.85, 436.85, 437.02, 437.43, 438.18, 438.18, 438.18, 438.68, 439.18, 477.27, 477.6, 531.93, 531.93, 533.6, 533.6, 616.27, 616.77, 617.43, 617.43, 617.93, 620.27, 620.43, 621.6, 621.93, 624.6, 647.85, 647.85, 647.85, 647.85, 650.35, 665.27, 666.27, 695.6, 695.6, 695.6, 712.43, 712.43, 712.43, 712.43, 713.1, 740.97, 744.97, 746.97, 754.8, 759.63, 759.63, 761.63, 801.13, 851.97, 856.97, 871.13, 871.47, 931.3
13	125.5, 153, 155.5, 157.5, 158, 176.5, 182.83, 191.83, 192.17, 193.17, 193.58, 198.33, 201.5, 217.83, 218.5, 218.92, 219.25, 219.58, 219.58, 219.92, 220.75, 223, 224.08, 225.17, 226.67, 227, 227, 228.67, 230, 231, 233.83, 233.83, 235.83, 241.33, 245.67, 255.67, 256, 256.5, 278.33, 278.5, 283.33, 283.67, 290.83, 293.33, 293.33, 293.5, 294, 294, 294.33, 294.33, 294.83, 294.83, 305.33, 306, 306.17, 314.17, 314.17, 316.83, 331.33, 332.17, 332.5, 333.5, 333.5, 334.5, 335, 335.5, 338.5, 342.5, 355.5, 356.5, 357, 372, 376, 378, 381, 381

continued ...

...continued

장비ID	고장시점(Hr)
14	48.5, 48.5, 53, 85, 104, 105, 124.3, 125.7, 129, 140, 158.17, 159.5, 160.67, 161.83, 161.83, 161.83, 161.83, 161.83, 186.67, 207, 213.25, 214.5, 215.17, 215.17, 215.92, 216.92, 221, 221.67, 222.58, 226.33, 226.33, 226.33, 241.67, 252.75, 255.08, 255.67, 257, 266.58, 266.58, 266.92, 266.92, 267.83, 283, 283, 283.67, 291.83, 312.25, 312.92, 313.58, 314.08, 314.58, 316.67, 317.33, 318, 318, 318.33, 318.5, 319.5, 319.5, 332.08, 335.08, 336.58, 348.92, 349.92, 350.17, 350.42, 350.42, 351.08, 356.83, 357.17, 366.33, 369.58, 371.33, 371.58, 388.75, 422.92, 423.58, 423.58
15	49, 49, 57, 80.5, 117.3, 125, 136.4, 136.4, 147.5, 148.17, 148.17, 149, 150.17, 150.5, 150.5, 150.5, 154, 194.33, 194.33, 194.67, 195.42, 195.42, 196.17, 196.75, 197.75, 199.67, 199.67, 199.67, 203.67, 207.75, 232, 233.17, 233.67, 234.83, 235.33, 236.83, 250.42, 272.92, 284.92, 285.75, 311, 311.17, 311.17, 312.33, 312.33, 312.5, 312.83, 313, 313.33, 313.67, 314.83, 315.5, 315.83, 316, 319, 319.17, 334.75, 338.75, 340.08, 353.83, 356.67, 356.67, 357.17, 365, 371.42, 371.75, 376.25, 376.58, 378.25, 378.42, 394.75, 395.75, 410.58, 410.58
16	92.5, 107.5, 110.5, 148, 149.67, 152.5, 153.33, 156.67, 158.67, 191.58, 191.58, 196.08, 205.42, 205.67, 206.58, 207.75, 210.67, 210.67, 212.67, 237.08, 238.25, 240.83, 245.92, 260.17, 278.5, 309.92, 311.92, 313.42, 314.25, 314.25, 314.75, 315.58, 316.58, 317.08, 318.08, 318.08, 333.92, 347.25, 348.92, 348.92, 349.42, 351.67, 357.33, 357.67, 362.67, 366.08, 369.42, 427.33, 428.5, 441.92, 441.92, 441.92

continued ...



...continued

장비ID	고장시점(Hr)
17	38.2, 38.2, 38.2, 47.87, 52.87, 102.12, 106.32, 119.12, 119.22, 130.62, 141.62, 141.62, 141.62, 162.58, 162.92, 165.42, 166.92, 166.92, 166.92, 166.92, 167.92, 169.42, 170.08, 202.92, 213.42, 214.33, 215.58, 217.48, 220.42, 221.58, 222.25, 223.42, 223.42, 245, 245.5, 246.5, 248.08, 252.88, 264.55, 265.47, 265.47, 287.38, 287.88, 301.38, 319.55, 321.38, 322.55, 322.72, 324.22, 324.22, 325.22, 327.55, 327.55, 329.38, 329.88, 329.88, 329.88, 343.88, 343.88, 358.05, 358.05, 359.72, 359.72, 359.72, 360.22, 362.88, 363.8, 364.13, 367.63, 367.97, 370.38, 372.05, 377.22, 378.88, 378.97, 391.55, 399.05, 402.38, 403.22
18	92.33, 133.93, 139.83, 168.63, 169.47, 169.8, 170.13, 172.8, 173.63, 173.63, 174.47, 174.47, 174.47, 174.47, 176.47, 176.47, 176.97, 179.97, 199.8, 199.8, 199.97, 204.47, 204.47, 214.05, 214.05, 214.3, 214.8, 216.22, 218.47, 218.47, 219.13, 219.13, 219.72, 219.72, 221.47, 221.63, 233.13, 233.13, 237.72, 245.05, 252.05, 262.55, 263.38, 263.72, 285.22, 285.72, 293.55, 293.88, 311.47, 312.97, 313.47, 314.3, 314.3, 314.88, 316.3, 317.13, 317.13, 323.97, 323.97, 338.8, 354.8, 354.8, 356.3, 356.3, 358.13, 360.3, 363.13, 363.13, 364.63, 366.3, 370.38, 370.9, 371.22, 371.3, 387.47, 388.63, 393.47, 397.8, 397.8
19	94, 110, 124, 138.02, 139.52, 141.52, 156.52, 201.68, 201.68, 214.68, 215.02, 215.35, 215.68, 218.68, 222.68, 225.68, 254.18, 300.68, 300.68, 301.68, 303.18, 303.68, 314.68, 318.18, 318.18, 321.18, 352.5, 352.5, 358.5, 365.5, 366.5, 369.5, 431.5, 440, 448.5
20	102.17, 136.17, 148.17, 148.17, 209.67, 249.33, 249.83, 260.83, 263.33, 288.83, 310.83, 338, 369.5, 373, 376, 391, 394, 401, 407, 412, 464, 470, 471
21	96.33, 111.33, 168.88, 169.88, 185.05, 224.55, 257.55, 260.05, 295.05, 295.38, 295.38, 299.55, 321.05, 352.83, 355.33, 395.33, 401.83, 403.83, 432.83, 432.83, 436.83, 482.97, 498.97, 511.97, 513.97, 521.97, 526.97

continued ...

...continued

장비ID	고장시점(Hr)
22	44.33, 89.67, 89.67, 104.67, 139.17, 139.17, 157.67, 173.5, 212.17, 224.33, 226.17, 226.83, 251.33, 255, 271, 275, 296.5, 299.5, 302.5, 303, 327.5, 334, 336, 356, 360, 406, 441.5, 442, 446
23	48, 103, 119, 153, 155, 155, 171.83, 197.83, 197.83, 213.33, 214, 244.33, 248.33, 261.33, 262.83, 297.83, 324.83, 331.33, 365, 367, 371, 378, 427, 461, 461
24	99, 116, 148, 148, 197.67, 221.67, 231.67, 261.67, 261.67, 265.17, 282.67, 285.67, 322.17, 334.67, 335.67, 360.67, 367.17, 370, 398, 407, 410, 411, 454, 480, 481
25	94.25, 148.25, 148.25, 161.75, 190.75, 204.25, 221.25, 222.75, 225.75, 267.75, 292.25, 293.25, 293.25, 293.75, 296.75, 298.25, 299.75, 324.75, 328.25, 331.25, 368.25, 372.25, 372.25, 377.25, 377.25, 379.25, 417.25, 426.25
26	42.5, 104, 135.5, 148, 149, 149, 163.83, 204.83, 204.83, 205.33, 218.33, 219, 221.33, 281.33, 311.83, 318.33, 321.83, 322.33, 322.33, 345.33, 351.83, 392, 403, 409, 457, 477, 486.5, 490, 492
27	31.33, 94, 123.5, 139, 140, 152.17, 194.83, 212.83, 219.33, 242.33, 242.33, 242.67, 242.67, 242.67, 242.67, 261.33, 261.33, 263.83, 296.83, 297.83, 308.33, 310.33, 311.33, 314.33, 330.33, 347.33, 350.83, 350.83, 390, 390, 394, 400, 402, 403, 442, 454
28	101.67, 149.67, 201.67, 219.17, 219.83, 221.67, 263.67, 267.17, 299.67, 303.17, 305.17, 306.17, 334.17, 370.67, 374.67, 379.67, 379.67, 380.67, 422.67, 423.67, 450.67, 455.67
29	29.5, 116.5, 119, 133.5, 143, 175.25, 175.42, 177.08, 178.08, 210.58, 226.58, 248.25, 251.08, 251.75, 251.92, 252.25, 252.42, 329.58, 335.58, 343.58, 358.58, 364.58, 368.58, 373.58, 373.58
30	33.5, 57.5, 83, 130, 130, 131, 145.33, 148.83, 156.83, 166.83, 167.98, 167.98, 212.25, 212.75, 225.25, 228.25, 232.92, 263.25, 263.25, 279.25, 285.25, 285.25, 307.25, 309.08, 309.75, 311.75, 312.08, 312.42, 313.42, 313.58, 340.25, 358.25, 361.25, 362.25, 362.25, 379.25, 379.25, 384.25, 384.25, 399.25, 400.25, 413.25, 415.25

continued ...

...continued

장비ID	고장시점(Hr)
31	121.5, 123, 137, 141, 152.15, 152.15, 199.25, 214.75, 217.75, 222.25, 223.25, 230.75, 257.08, 257.75, 297.75, 299.25, 301.42, 301.75, 302.58, 302.58, 321.75, 337.75, 339.75, 342.08, 343.75, 357.75, 368.75, 382.75, 384.75, 398.75, 400.75
32	31.5, 55.5, 134, 152.4, 203, 203.5, 213, 215, 218.67, 225.5, 246.33, 246.67, 254, 259, 263.67, 284.5, 286.5, 286.67, 287.17, 287.17, 287.5, 287.83, 307, 323, 326, 346, 351, 366.75, 369, 375
33	126, 138.5, 164.67, 164.67, 172.83, 177.83, 205.33, 217.83, 217.83, 226.83, 250, 266.5, 267, 291.5, 292.17, 295.5, 296, 296.5, 296.5, 296.5, 296.67, 308.83, 321.33, 341.33, 342.33, 343.33, 348.33, 377.33, 403.33, 409.33
34	116.5, 117, 130.17, 133, 141.5, 141.75, 156.25, 210.25, 215.25, 222.25, 245.58, 246.25, 258.25, 265.58, 286.25, 287.75, 290.75, 290.92, 291.25, 291.58, 291.92, 291.92, 292.08, 304.25, 311.25, 317.25, 331.25, 333.25, 335.25, 337.58, 339.25, 360.25, 376.25, 376.25, 387.25, 400.25, 411.25
35	55, 55, 86, 117, 117, 117, 119.5, 130.42, 156.25, 202.25, 203.25, 217.75, 219.75, 223.75, 225.42, 256.92, 269.25, 276.75, 302.25, 303.75, 307.25, 308.75, 309.25, 309.42, 310.42, 310.58, 314.75, 335.25, 342.25, 344.25, 358.25, 360.25, 363.42, 364.25, 384.25, 384.25, 390.25, 394.25, 396.25, 413.25, 417.25, 425.25
36	33.5, 46.5, 90.5, 109.5, 122, 141, 158.4, 210.67, 211.67, 223.17, 227, 263.83, 264.5, 264.5, 282.5, 291, 318.17, 323.83, 324.83, 325.5, 350.17, 367.17, 370.17, 373.17, 375.17, 380.17, 413.17, 422.17, 433.67
37	34, 79.5, 109.5, 119, 133, 158.4, 212, 234.5, 235.17, 238.83, 275.67, 292.33, 300.67, 302.33, 326.33, 331.33, 332, 332.5, 333, 333, 333.17, 376.83, 377.83, 380.83, 418.83, 436.83, 450.83
38	72.5, 123, 151.5, 210.5, 212, 223.5, 229.83, 240.33, 241.83, 282.83, 321.67, 323, 325, 325.17, 325.83, 344.5, 363.5, 365.5, 365.5, 383.5, 413.5, 419.83

continued ...

...continued

장비ID	고장시점(Hr)
39	94, 111.43, 111.43, 111.43, 111.43, 111.97, 143.85, 180.65, 193.68, 204.98, 305.88, 319.15, 322.32, 326.98, 326.98, 326.98, 328.15
40	74, 77, 78.67, 93, 93, 147.92, 188.42, 191.75, 192.08, 192.58, 194.17, 200.58, 202.42, 204.83, 231.92, 239.83, 253.17, 253.17, 256.17, 257, 257, 280.83, 282.33, 289.17, 297.5, 335.33, 338.17
41	74.05, 81.58, 86.58, 96.15, 97.48, 97.48, 98.73, 98.9, 157.23, 187.77, 190.1, 190.1, 190.85, 329.08, 333.75, 333.75, 337.25, 355.45, 355.95
42	72, 73, 91.97, 94.72, 95.88, 215.38, 216.38, 230.8, 269.97, 271.8, 310.05, 328.35, 334.85, 388.52
43	131, 160, 178.5, 183.5, 187.5, 190.5, 205, 205.5, 205.5, 205.5, 225, 249.5, 253.5, 261.5, 265, 266.5, 266.5, 283.5, 285.5, 288.5, 288.5, 291.5, 293.5, 345, 345, 352.5
44	141, 162, 169, 169, 170, 179, 194.5, 224, 228, 231, 232, 234, 237, 240, 266, 268, 295, 303.5
45	174.37, 174.37, 174.37, 184.2, 203.03, 204.03, 231.53, 238.53, 239.53, 239.53, 240.53, 240.53, 267.03, 274.53, 299.53, 303.53, 303.53
46	71.03, 153.53, 172.53, 173.53, 184.7, 185.2, 185.2, 185.2, 185.7, 190.2, 203.2, 203.2, 231.2, 240.2, 241.2, 243.2, 255.2, 276.2, 281.2, 283.2, 333.2, 333.2, 341.2
47	135, 164.17, 172.17, 172.17, 175.67, 179.67, 202.17, 202.17, 228.17, 231.17, 236.17, 238.17, 240.17, 240.17, 273.17, 274.17, 274.17, 276.17, 284.17, 312.67, 312.67
48	60, 64, 72, 145.02, 146.52, 146.52, 147.18, 150.68, 150.68, 162.18, 179.02, 179.52, 179.52, 181.52, 189.02, 189.02, 206.02, 206.02, 235.52, 237.52, 237.52, 237.52, 238.52, 239.52, 240.52, 245.52, 252.52, 285.52, 287.52, 287.52, 290.52, 323.02, 323.02, 349.02
49	59.5, 168, 168.5, 175.5, 175.5, 192.5, 195.5, 196.5, 197.5, 198, 198, 198.5, 217.5, 218, 218.5, 220, 220, 222, 228.5, 256, 258, 258.5, 259, 277.5

continued ...

...continued

장비ID	고장시점(Hr)
50	147.5, 167.5, 168, 169.5, 170.5, 177.5, 178.5, 178.5, 179, 179, 197, 198, 198, 201.5, 201.5, 201.5, 224, 225.5, 225.5, 226, 226.5, 227, 227.5, 229, 232, 233, 255, 257, 258, 258.5
51	87, 87, 99.5, 173.5, 175, 176, 176, 177.5, 177.5, 177.5, 178, 196, 197, 200.5, 201, 223, 224, 224.5, 225, 226, 226, 226.5, 227.5, 229.5, 230.5, 255.5, 257, 272, 277, 295.5
52	173.5, 173.5, 175, 183, 184, 184, 184.5, 203, 204, 206.5, 206.5, 206.5, 207, 230, 231, 231.5, 231.5, 235, 236, 243, 243, 274, 275, 275.5, 277.5, 323.5
53	83, 83, 83, 83, 159.5, 169, 188.5, 190, 192.5, 193, 215, 216, 220, 225.5, 248, 250.5, 271
54	78, 87, 164.5, 165.5, 172, 172, 172.5, 172.5, 172.5, 172.5, 173, 191, 191.5, 191.5, 194, 194.5, 222.5, 231.5, 234.5, 261.5, 262.5, 263.5, 295.5
55	170, 171, 180, 180, 203, 203, 205.5, 205.5, 228, 231, 231, 236.5, 268, 273.5, 290, 298, 306.5
56	91.5, 166.5, 179, 179, 198, 199, 199, 202.5, 202.5, 203, 226, 226.5, 227, 227, 227.5, 228.5, 230, 232.5, 233, 233, 260, 282, 291.5, 301

# **Abstract**

## **A reliability trend analysis for heterogeneous repairable systems with left truncated data**

Ilyong Na

Financial Risk Engineering Lab.

School of Industrial Engineering

The Graduate School

Seoul National University

Keeping the quality in use of high-tech systems steady is difficult. Because of the complexity of the product, the likelihood of failure is high, while detection and recovery are difficult. Therefore, product design and operation maintenance policy that minimizes the occurrence of failures and restores the failures in the shortest time with minimum effort is considered to be the most important factor. The most influential factor in this activity is “reliability”. The reliability of lexical meaning is defined as “probability that can be used without failure for a fixed time in a given condition” and the key keyword is “failure”.

To minimize failures, developers often use a variety of methods, such as using highly reliable components, duplicating parts, and using simplified components. If failures can not be minimized, maintenance activity can

be made easier and shorter recovery times. However, after the completion of the development, it is difficult to make further improvements, so focusing on activities to maintain the status rather than improvement. That is, by monitoring the equipment status, you observe whether the “failure trend” is increasing, decreasing, or maintaining a certain level. In a repairable system, the Power Law model is widely used as a model to observe the trend of failure. The Power Law model is an NHPP model using the Weibull type intensity function. Due to the simplicity of the model and the ease of use, the Power Law model has been used continuously in the fields related to the repairable system such as the weapon system since it was proposed by Crow in the 1970s. Despite this simplicity and convenience, there is a constraint that the use of the model requires the use of steady data collected from the beginning of the operation of the equipment. In particular, military equipment with a life cycle of more than 10 years, in contrast to commercial products that can easily track and maintain fault data with a short life cycle, may have limited use of the model due to data quality issues. In Korea, data loss frequently occurs due to frequent change of data management personnel or recording system during the life cycle of equipment. Therefore, since the amount of data is large, most of the data is cut off to the left, which limits practical analysis models. In particular, it is difficult to use for analysis such as reliability trend analysis.

The purpose of this study is to present a practical model of the power law model that can be used to analyse the reliability trends. In order to reflect the left truncated condition to the power law model, we propose the use of the E-M algorithm to compensate for the loss. The E-M algorithm is a

general method to compensate for data loss, and is an optimal method when using ML (Maximum Likelihood) for parameter estimation of probability models as in this paper. However, there is a disadvantage that the global optimization can not be guaranteed. To overcome this problem, we propose a method of initializing the parameter estimates by empirical methods. In addition, even if the same equipment is used, the extended model can be considered because it can show different failure trends depending on the usage method, the user, and the usage environment. The former is called the “Basic model” and the latter is called “Extended model”. The extended model suggested that the observable heterogeneity factor, Covariates and Frailty, can be considered at the same time.

In order to verify the utility and range of the proposed model, various data loss conditions, power law parameters and Frailty level combinations were simulated, and the range and error levels of the proposed model were confirmed. In the case of the basic model, it was confirmed to be effective at data loss level of 10 % or less. In addition, it has been confirmed that the extended model shows a more uniform estimation error than the basic model regardless of the data loss level. Finally, by applying the proposed two models to the field failure data of the tanks in use in Korea, the reliability level of the equipment is calculated, and a simple maintenance policy decision making example using it is presented, so that it can be fully utilized for the actual maintenance policy decision.

This study started from a problem consciousness to effectively deal with the left truncated data, and could provide a model that yields valid results within a certain range. However, frequent use of large sample assump-



tions, limited presentation of mathematical models, and sensitivity analysis within a limited range should be gradually complemented by further studies.

**Keywords :** Left truncated, Reliability trend, Heterogeneity, Frailty, Co-variates, Basic Model, Extended Model

**Student Number :** 2011-30312